(ii) Rejections Under 35 U.S.C. §102. Claims 1, 6-8 and 26-30 remain rejected as allegedly anticipated by Aghajari et al., Protein Science, 1998, 7:564-572 ("Aghajari"). The rejection is traversed on the grounds that Aghajari does not disclose the instantly claimed methods.

In first response to the rejection, the claims have been amended to be directed to methods that call for "utilizing the calculated dimensions, chemical and/or electrostatic properties obtained in step b [or steps b and c] to identify compounds that contain functional groups that can be accommodated by said cavity." *See* step "c" in claims 1, 26, 27, 29 and 30 and step "d" of claim 28. Aghajari merely defines the coordinates of a chloride ion that is already present in the allosteric site of the α-amylase enzyme. Thus, Aghajari fails to utilize any information concerning the physical properties of a cleft to identify compounds that contain functional groups that can be accommodated by a cavity. For at least this reason, Aghajari does not anticipate the instant claims.

On at a more basic level, the disclosure of Aghajari is limited to the description of the binding of a chloride ion on an α-amylase enzyme. A chloride ion is not a "compound." Neither does a chloride ion contain a "functional group." Accordingly, Aghajari fails to identify a compound that contains functional groups that can be accommodated by a cavity on a protein. For at least this reason, additionally, Aghajari does not disclose the instantly claimed methods.

Finally, Aghajari does not anticipate the instant claims because Aghajari does not explicitly or inherently disclose a "method of identifying a compound that modulates interaction between a functionally critical site of a target protein and a modifier" (as recited in the preamble of e.g., claim 1) comprising the steps set forth in the claims and "*thereby identifying* said compound that modulates intermolecular interaction at said functionally critical site of said target protein and *said* modifier," as recited at the end of each of the independent claims (emphasis added). Aghajari merely describes the crystal structure of α-amylase with a bound activating chloride ion. Aghajari does not use the crystal structure to identify the chloride ion as a modulator of α-amylase function. Indeed, this activity was previously known. Thus, by no fair interpretation of the phrases "method of identifying a compound" and "thereby identifying said compound" that are recited respectively in the preamble and wherefore-clause of the claims can Aghajari be considered to anticipate the instant claims

For at least all of the reasons set forth above, Aghajari does not anticipate the instantly claimed methods. Reconsideration of the instant claims and withdrawal of all rejections thereof as anticipated by Aghajari is request, accordingly.

CONCLUSION

This application is believed to be in condition for allowance, which is earnestly solicited.

If there are any other issues remaining which the Examiner believes could be resolved through either a Supplemental Response or an Examiner's Amendment, the Examiner is respectfully requested to contact Applicant's representative at the telephone number indicated below.

Dated: September 25, 2006

Respectfully submitted,

By 

Mitchell Bernstein, Ph.D.
Registration No.: 46,550
DARBY & DARBY P.C.
P.O. Box 5257
New York, New York 10150-5257
(212) 527-7700
(212) 527-7701 (Fax)
Attorneys/Agents For Applicant

J. Mol. Biol. (1995) 254, 771–792

JMB

Molecular Dynamics Simulation of *E. coli* Ribonuclease H_I in Solution: Correlation with NMR and X-ray Data and Insights into Biological Function

Marios Philippopoulos[1] and Carmay Lim[2]*

[1]*Department of Chemistry University of Toronto Toronto, Ontario M5S 1A1 Canada*

[2]*Institute of Biomedical Sciences, Academia Sinica and Department of Chemistry, National Tsing-Hua University Taipei Taiwan 11529 Republic of China*

**Corresponding author*

A 500 ps molecular dynamics simulation of *Escherichia coli* RNase H_I in the presence of explicit water molecules has been carried out to aid in the interpretation of NMR N–H backbone model free parameters and X-ray *B*-factor values of the free enzyme. Both experimental techniques have revealed unusual structural and dynamic features of the protein. Atomic fluctuations (*B*-factors) and re-orientational motions of the backbone heteronuclear bonds (order parameters) computed from the simulation are compared with results obtained from experiments. Qualitative agreement is obtained between the computed and X-ray *B*-factors, whereas the agreement between the computed and NMR generalized order parameters is as good as quantitative for most residues. Reasons for significant discrepancies, the physical basis and the plausible biological consequences of the observed protein dynamics are discussed.

Keywords: RNase H; molecular dynamics; order parameters; *B*-factors; solution structure

Introduction

Ribonuclease H (RNase H) is an endonuclease that specifically cleaves the RNA strand of a DNA·RNA hybrid, yielding a 3'-hydroxyl and a 5'-phosphate. Experimental data suggest that (1) RNase H of cellular origin is involved in DNA replication or repair. (2) RNase H of retroviral origin is required to reverse-transcribe the viral single-stranded RNA genome into double-stranded DNA (Crouch, 1990). RNase H cleavages at the origin of plus strand DNA synthesis are species specific (Luo *et al.*, 1990). Because of its crucial role in infection by retroviruses, RNase H is a potential target for drugs against retroviral diseases, like HIV infections and AIDS.

Among the family of RNase H domains, *Escherichia coli* RNase H_I is the most extensively studied. It consists of 155 amino acid residues containing five α helices and five β strands, as well as five reverse turns (Katayanagi *et al.*, 1992), three of which are part of two loops and the C-terminal region. The secondary structure elements are listed in Table 1 according to the definitions employed by Yang *et al.* (1990; see Figure 1). The *E. coli* sequence numbering is employed in this paper. RNase H_I has an isoelectric point of 9.0 (Kanaya *et al.*, 1989) and an optimal pH of 8.0 for enzymatic activity (Oda *et al.*, 1994). There are seven highly conserved residues in the enzyme (Yang *et al.*, 1990); Asp10, Glu48, Asp70, Ser71, His124, Asn130 and Asp134 (see Figure 1). The first three form the catalytic active site. The residues responsible for substrate binding and catalytic reaction have been identified from the following. (1) Site-directed mutagenesis (Kanaya *et al.*, 1990, 1991a,b) and chemical modification studies (Hogrefe *et al.*, 1990; Uchiyama *et al.*, 1994). (2) X-ray structures of *E. coli* RNase H_I solved by multi-wavelength anomalous diffraction to 2.0 Å (Yang *et al.*, 1990; protein data base (PDB) entry 1RNH) and multiple isomorphous replacement to 1.48 Å (Katayanagi *et al.*, 1992; PDB entry 2RN2), a 2.8 Å X-ray structure of *E. coli* RNase H_I complexed with a metal ion (Katayanagi *et al.*, 1993a), crystal structures of active site mutants of *E. coli* RNase H_I (Katayanagi *et al.*, 1993b), and model structures of the enzyme-substrate (Yang *et al.*, 1990; Nakamura *et al.*, 1991; Katayanagi *et al.*, 1992). (3) NMR studies of nucleic acids and metal ions binding to *E. coli* RNase H_I (Nakamura *et al.*, 1991; Oda *et al.*, 1991, 1993b, 1994; Huang *et al.*, 1994). pK_a values of histidines (Oda *et al.*, 1993a)

Abbreviations used: MD, molecular dynamics; rmsd, root-mean-square deviation; NOE, nuclear Overhauser enhancement; IMAC, immobilized-metal-affinity-chromatography; PDB, protein data base.

and carboxyl groups (Oda *et al.*, 1994); and backbone dynamics of the free enzyme (Mandel *et al.*, 1995). The results from these studies are summarized below.

Binding/specificity

The loops between β_1 and β_2, between α_C and α_D (the "handle" region), and between β_5 and α_E have been implicated in substrate binding (Mandel *et al.*, 1995). The crystal structures of Mg^{2+}-free (Yang *et al.*, 1990; Katayanagi *et al.*, 1992) and Mg^{2+}-bound (Katayanagi *et al.*, 1993a) RNase H_I indicate; (1) at least one metal cation binding site involving the side-chains of Asn44 and Glu48. (2) A phosphate binding site roughly 13.7 Å from the catalytic site at the NH_2 termini of α_A and α_D. Site-directed mutagenesis and NMR titration experiments indicate that Cys13, Asn16, Asn44, Asn45 and Gln72 are engaged in hybrid binding (Nakamura *et al.*, 1991), in addition to the active site acidic residues, Asp10, Glu48 and Asp70 (Kanaya *et al.*, 1990). Furthermore, the NMR-observed enhanced backbone dynamics of residues 89 to 99 in the handle region, Gly15 and Asn16 in the β_1–β_2 loop and Gly123, Ala125 and Gly126 in the β_5–α_E loop suggest a role in conforming the substrate-binding cleft to the substrate surface and positioning catalytic residues (Mandel *et al.*, 1995).

Catalysis

The active site residues Asp10, Glu48 and Asp70 are crucial for activity: amidation of any of these carboxylates (Asp → Asn and Glu → Gln) abolished activity, and single methylene shifts in side-chain length (Asp → Glu and Glu → Asp) reduced k_{cat} to 1 to 8% that of the wild-type without a change in K_m (Kanaya *et al.*, 1990). Although Asp134 is not crucial for activity (as the D134H and D134N mutant proteins retained considerable activity), it has been suggested that it takes part in catalysis but not in binding, since the replacement of Asp134 by residues other than asparagine and histidine dramatically decreased the enzymatic activity without seriously affecting substrate binding and tertiary structure (Haruki *et al.*, 1994). His124 is also involved in the catalytic function, as its mutation to Lys, Gln or Glu dramatically lowered the k_{cat} value, without significantly affecting the K_m value, as compared to those of the native enzyme (Oda *et al.*, 1993a). Two mechanisms have been proposed for the catalytic function of the enzyme. (1) A single-metal-ion carboxylate-hydroxyl relay mechanism (Nakamura *et al.*, 1991). (2) A two-metal-ion mechanism first suggested for *E. coli* DNA polymerase I (Yang *et al.*, 1990; Beese & Steitz, 1991). The latter is supported by the finding that Mn^{2+} binds at two sites in the RNase H domain of HIV reverse transcriptase (Davies *et al.*, 1991), whereas the former is supported by the identification of a single divalent cation binding site in *E. coli* RNase H_I by X-ray (Katayanagi *et al.*, 1993a), NMR (Oda *et al.*, 1991), chemical modification and kinetic studies (Uchiyama *et al.*, 1994).

As both NMR spectroscopy and X-ray crystallography have revealed unusual structural and dynamical features of *E. coli* RNase H_I, we have carried out a 500 ps molecular dynamics (MD) simulation of the enzyme in the presence of explicit water molecules. There are five main objectives of our study. First, to evaluate the ability of MD simulations to reproduce dynamic behaviour on the ps timescale. Second, to complement NMR spectroscopic and X-ray crystallographic data, e.g. by

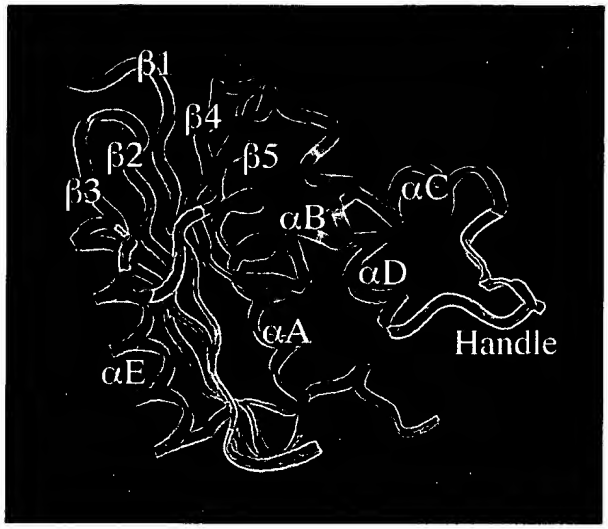

Figure 1. Ribbon diagram of the backbone trace of *E. coli* RNase H_I; conserved residues and regions involved in substrate binding are colored red and yellow, respectively.

Table 1. Average *B*-factors for secondary structure elements

Secondary structure	Residue range	*B*-factor (Å2) MD	2RN2	1RNH
β1	4–13	5.5	14.8	12.3
β2	18–27	5.2	13.9	10.1
β3	32–42	6.6	15.8	13.4
β4	64–69	6.0	14.4	13.5
β5	115–120	8.1	16.8	16.5
αA	43–58	5.8	12.5	6.7
αB	71–80	7.2	16.6	8.2
αC	81–88	10.3	25.2	9.8
αD	100–112	6.7	13.5	9.9
αE	127–142	6.7	17.4	15.0
$Turn_{12}$	14–17	10.2	23.3	18.3
$Turn_{23}$	28–31	25.2	31.0	34.7
$Loop_{A4}$	59–63	15.8	23.0	26.7
$Handle_{CD}$	89–99	22.7[a]	18.0	9.0
$Loop_{5E}$	121–126	17.9	28.0	25.9
N Terminus	1–3	109.7[a]	28.7	40.5[b]
C Terminus	143–155	30.1[a]	28.9	16.6[b]

MD denotes *B*-factor values calculated from the 400 ps production trajectory; 2RN2 and 1RNH refer to *B*-factor values taken from Katayanagi *et al.* (1992) and Yang *et al.* (1990), respectively.

[a] Certain residues in this element exhibit unrealistically high fluctuations, probably due to inadequate solvation.

[b] Residues 1 and 153 to 155 are disordered in the 1RNH structure.

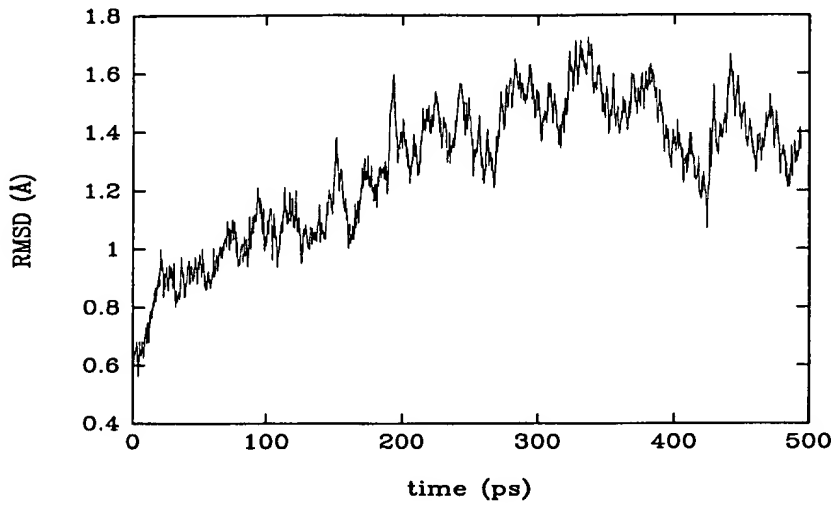

Figure 2. Plot of the root mean square deviation (RMSD) of the backbone from the 2RN2 crystal structure as a function of simulation time.

determining order parameters of catalytically important residues which could not be quantified by NMR (Mandel *et al.*, 1995) and evaluating conformational changes of residues involved in crystal contacts or disordered in the X-ray structure. Third, to assess the degree of overlap between NMR spectroscopy and X-ray crystallography in describing protein mobility. Fourth, to provide a physical basis for the protein dynamics observed by NMR spectroscopy (Mandel *et al.*, 1995) and X-ray crystallography (Yang *et al.*, 1990; Katayanagi *et al.*, 1992). Fifth, to correlate the structural and dynamical features of certain residues/regions to biological function. Thus, both atomic fluctuations (*B*-factors) and re-orientational motions of heteronuclear bonds (order parameters) computed from the simulation are compared with experiment. The reasons for significant discrepancies, the physical basis and the plausible functional consequences of the observed protein dynamics are presented in the Discussion section.

Results

A plot of the root-mean-square deviation (rmsd) of the backbone from the 2RN2 structure as a function of simulation time (Figure 2) shows that the rmsd initially rises but after 100 ps it fluctuates between 1.0 and 1.7 Å. These values are consistent with those found in other protein simulations. Thus, the first 100 ps is included in the equilibration phase, whereas the remaining 100 to 500 ps part of the trajectory is employed in analysis. Configurations from every 0.1 ps of the production trajectory have been averaged to yield a mean structure. In computing the *B*-factors (Brooks *et al.*, 1988), order parameters (Levy *et al.*, 1981, 1982) and effective internal correlation times (Philippopoulos & Lim, 1994), overall translational and rotational diffusion have been removed by superimposing the backbone of each configuration in the 400 ps trajectory onto the backbone of the 2RN2 structure using a mass-weighted least-squares fitting algorithm (Kabsch, 1976). A list of hydrogen bonds involving conserved residues that are observed in the simulation is presented in Table 2, where Wat denotes a crystal water molecule whose numbering is taken from Katayanagi *et al.* (1992), and Bul denotes water overlaid in the simulation. A hydrogen bond is defined by a heavy-heavy and a

Table 2. Hydrogen bonds of conserved residues (and the backbone carbonyl oxygen of the preceding residue) observed in the simulation

Donor/ Acceptor	Donor/ Acceptor	r_{Xray} (Å)	$\langle r \rangle_{MD}$ (Å)	rmsd (Å)
Thr9(O)	Wat175(O)	2.56	2.71	0.12
Asp10(N)	Gly23(O)	3.02	2.89	0.14
Asp10(O)	Gly23(N)	2.98	2.94	0.16
Asp10($O^{\delta 2}$)	Bul1184(O)	—	3.42	2.02
Met47(O)	Ala51(N)	3.08	3.14	0.21
Glu48(N)	Asn44(O)	3.13	2.97	0.15
Glu48(O)	Bul3341(O)	—	2.79	0.16
Glu48($O^{\epsilon 1}$)	Ser71($O^{\gamma 1}$)	2.60	2.71	0.13
Glu48($O^{\epsilon 1}$)	Bul4232(O)	—	2.86	0.59
Glu48($O^{\epsilon 2}$)	Wat156(O)	3.32	2.83	0.31
Glu48($O^{\epsilon 2}$)	Bul3158(O)	—	3.56	1.70
Thr69(O)	Val121(N)	3.11	3.08	0.22
Asp70(N)	Asp70($O^{\delta 1}$)	2.87	2.87	0.21
Asp70($O^{\delta 1}$)	Wat166(O)	2.82	2.94	0.50
Asp70($O^{\delta 1}$)	Wat166(O)	2.78	2.93	0.33
Asp70($O^{\delta 1}$)	Wat175(O)	3.39	2.70	0.15
Asp70($O^{\delta 2}$)	Wat166(O)	2.95	3.10	0.38
Asp70($O^{\delta 2}$)	Bul3158(O)	—	2.90	0.57
Ser71(N)	Thr69($O^{\gamma 1}$)	3.19	3.20	0.14
Ser71(O)	Arg75(N)	3.26	3.33	0.14
Ser71($O^{\gamma 1}$)	Val74(N)	2.90	3.22	0.20
Glu129(O)	Cys133(N)	3.08	2.91	0.14
Glu129(O)	Cys133(S^{γ})	3.42	3.23	0.16
Asn130(N)	His127(O)	2.88	3.12	0.21
Asn130($N^{\delta 2}$)	Wat175(O)	3.26	3.12	0.28
Cys133(O)	Ala137(N)	2.84	2.93	0.13
Asp134(N)	Asn130(O)	3.08	3.03	0.18
Asp134(O)	Arg138(N)	2.82	3.00	0.16
Asp134($O^{\delta 1}$)	Arg138(N^{ϵ})	3.16	3.23	0.60
Asp134($O^{\delta 1}$)	Bul1184(O)	—	3.27	1.86
Asp134($O^{\delta 2}$)	Arg138(N^{ϵ})	3.05	3.22	0.32

The r_{Xray} distances are from the 2RN2 structure after hydrogens have been built in; $\langle r \rangle_{MD}$ and rmsd are average distances from the production trajectory and their corresponding rmsd, respectively.

light-heavy donor-acceptor distance of $\leqslant 3.5$ Å and $\leqslant 2.5$ Å, respectively.

Structural analysis

Most protein–protein charge–charge interactions (involving side-chains) observed in the 2RN2 structure (Katayanagi *et al.*, 1992) are preserved in the simulation. In particular, the hydrogen bond network involving Arg46, considered responsible for rigidifying the overall protein conformation, has remained intact. A number of interactions between the C terminus (starting at Asn143) and residues from β3 and αA are maintained in the simulation, as well as a crystal water molecule bridging the C terminus and β3, Asp148(N)–Wat173–Arg41(N). These contacts serve to rigidify the C-terminal conformation (see also Discussion).

However, there are some protein–protein charge–charge interactions that are in dynamic exchange or are not preserved; these involve residues directly participating or in the immediate vicinity of residues involved in inter-molecular contacts in the 2RN2 structure (Katayanagi *et al.*, 1992). Therefore, some of the non-preserved interactions appear to be a result of the local environment relaxing to energetically more favorable conformations during the simulation. For example, in the 2RN2 structure, the backbone and side-chain oxygens of Asp94 have inter-molecular contacts with neighbouring Arg132 and Gln4 and this may have resulted in the disruption of the Asp94($O^{\delta 2}$)–Thr92($O^{\gamma 1}$) hydrogen bond in the handle region during the simulation. Furthermore, the side-chains of Arg29, Arg31, Asp70 and Arg132 hydrogen bond with residues in a neighbouring molecule in the 2RN2 structure; the close proximity of Arg29, Arg31, Asp70 and Arg132 to Arg27, Glu32, Ser68 and Asn130, respectively, may have caused several interactions between the β sheet and αE involving the latter residues to be unstable in the absence of crystal contacts: the Ser68($O^{\gamma 1}$)–Asn130($N^{\delta 2}$) interaction is partly preserved and two out of the four hydrogen bonds involving Arg27 in the 2RN2 structure are not stable in the simulation.

Another factor that may affect some of the changes in the hydrogen bonding patterns relative to the X-ray structure is the different pH states of the X-ray structure (pH 9.0) and the simulated one (pH 5.5), resulting in different protonation states for Asp10 and the histidines. For example, in the 2RN2 structure, Asp10($O^{\delta 2}$) is found to be 2.63 and 3.23 Å away from the backbone nitrogen and oxygen of Gly11, respectively; the proximity of an anionic Asp10($O^{\delta 2}$) to the carbonyl oxygen of Gly11 is, however, surprising and may be due to an inter-molecular non-bonded contact with neighbouring Lys87(N^{ζ}). In the dynamics-averaged structure, the neutral side-chain atom of Asp10($H^{\delta 2}$) hydrogen bonds instead to a water molecule. As another example, the reported intramolecular interactions in the 2RN2 structure between $N^{\delta 1}$ and $N^{\epsilon 2}$ of His114 and the backbone carbonyl oxygens of Ala110 and Glu61, respectively, is in apparent contradiction to the singly protonated state of His114 expected under the crystallization conditions (pH 9) and its measured pK_a value ($pK_a < 5$). These two hydrogen bonds were not found when hydrogen atoms were built into the 2RN2 structure and they are not seen in the average structure, as deprotonated His114($N^{\delta 1}$) and protonated His114($N^{\epsilon 2}$) are hydrogen-bonded to Cys63(H^{N}) and Leu59(O), respectively.

In contrast to protein–protein interactions, most protein–water hydrogen bonds found in the 2RN2 structure are not maintained during the simulation. Several of these non-preserved interactions are mainly between C-terminal residues and water molecules, as well as all hydrogen bonds of Glu147 and residues C-terminal to Tyr151 with water molecules. However, a number of water molecules still play a structural role in the dynamics average structure. In many cases, these water molecules bridge different secondary structure elements, thereby contributing to the overall tightening of the protein conformation (see Table 2).

Protein–protein and protein–water hydrogen bonds involving backbone amide nitrogens (and carbonyl oxygens by virtue of the peptide bond) are of particular interest in this study, as they provide insight into the physical basis of backbone dynamics observed by NMR. Most of the backbone–protein hydrogen bonds observed in the 2RN2 structure persist throughout the trajectory. The exceptions involve mostly hydrophobic residues in helices and loops and charged residues in the vicinity of hydrogen bonding contacts with another molecule in the 2RN2 lattice (e.g. Tyr28, Arg31 and Asp148) as well as His127, which undergoes a conformational change, probably due to its positively charged state in the simulation. In contrast, many backbone–crystal water hydrogen bonds are not preserved, although in several cases, the 2RN2 water molecule involved in the hydrogen bond is either replaced by another crystallographic or a bulk water molecule. Furthermore, several hydrogen bonds involving crystal water molecules occur in the dynamics average but not in the 2RN2 structure.

Isotropic *B*-factors

Both theoretical and experimental *B*-factors of the N, C^{α}, and C(=O) backbone atoms are averaged for each secondary structure element and they are listed in Table 1. The experimental values are taken from PDB entries 2RN2 (Katayanagi *et al.*, 1992) and 1RNH (Yang *et al.*, 1990), whereas the calculated values are based on the 100 to 500 ps portion of the trajectory. The average *B*-factors correlate with secondary structure. Residues in β-strands and α helices have generally lower mean-square displacements (*B*-factors) than residues in non-regular structures; the termini of the polypeptide chain have very high mean-square displacements, in

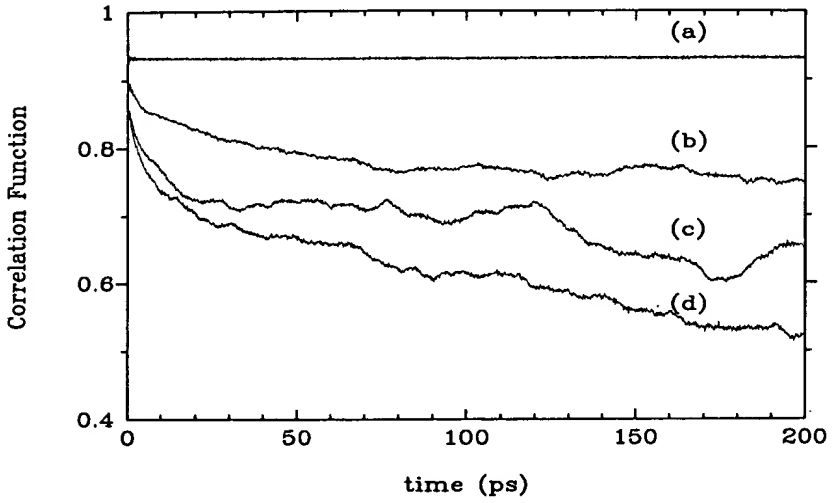

Figure 3. Representative re-orientational auto-correlation functions for the backbone N-H bond from each of the four classes described in the text. (a) Met50 (class I). (b) Lys96 (class II). (c) His124 (class III). (d) Gly123 (class IV).

accord with the finding that the electron densities for these regions are very weak in the 2RN2 structure and are absent in the 1RNH structure. Quantitative agreement (within the experimental uncertainty) is obtained between the 1RNH and computed average *B*-factors for the first three helices, αA, αB, and αC (Table 1). Apart from the terminal and handle regions, the computed average *B*-factors are generally lower than the X-ray values, as found in earlier studies (Petsko & Ringe, 1984; Post, 1989; Chandrasekhar *et al.*, 1992); possible reasons for this trend are presented in Discussion.

N–H Re-orientational auto-correlation functions

The re-orientational auto-correlation functions for the backbone N-H bond of 149 out of the 155 residues (i.e. excluding the N terminus and five prolines) have been calculated from the molecular dynamics trajectory over the period 100 to 500 ps. They can roughly be categorized into four classes. (1) Class I: correlation function decay within 20 ps of correlated time to a well defined plateau value. (2) Class II: decay to a plateau value within a correlated time of more than 20 ps. (3) Class III: initial decay to a plateau value, then further decay with no plateau reached. (4) Class IV: no plateau reached within 200 ps correlated time. Representative correlation functions from each class are shown in Figure 3. The correlation function of Gly126 is the most erratic: large oscillations are exhibited throughout the correlated time. Out of the 115 class I residues, 102 belong to regular secondary structure elements, i.e. α-helix or β-strand. In contrast, only five of the 34 residues whose correlation functions belong to classes II to IV (all of them in class II) belong to an α-helix or β-strand. In the handle region, residues Lys91, Thr92 and Asp94-Lys96 belong to class II, whereas Ala93 belongs to class III.

Generalized order parameters

The generalized order parameters S^2_{MD} of the 149 backbone N-H bonds calculated from the 100 to 500 ps portion of the molecular dynamics trajectory are listed in Table 3 and compared with corresponding experimental data (S^2_{NMR}) in Figure 4 (Mandel *et al.*, 1995). S^2_{MD} values and their uncertainties for N-H pairs with correlation functions that did not reach a well-defined plateau value within 200 ps correlated time are listed with a question mark beside the value in Table 3. For 48 out of 116 residues quantified experimentally, the S^2_{MD} differ by more than 0.055 from S^2_{NMR}; possible reasons for this discrepancy will be presented in Discussion. Discord with experimental values is most pronounced (greater than 0.2) for residues 2–3, 29 to 31, 71, 126 and 155 (Figure 4).

Generalized order parameters of the backbone C^α–H^α atom pairs of 141 residues (14 glycines excepted) have also been computed using equation (6) and they are compared to the N-H S^2_{MD} values in Figure 5. The C^α–H^α S^2_{MD} values are generally higher than corresponding N-H values throughout the primary sequence: the average differences for β strands and α helices are 0.061 and 0.040, respectively. (For only seven residues, Leu14, Arg27, Tyr28, Met50, Val54, Lys60 and Gln76, are the differences between the two sets of S^2_{MD} values smaller than 0.01.) The same trend has also been observed in molecular dynamics studies of a zinc finger peptide (Palmer & Case, 1992) and BPTI (Smith *et al.*, 1995). The restricted mobility of C^α–H^α relative to the N-H bond may reflect the fact that the peptide group can rotate relatively freely with little energetic cost to the rest of the backbone conformation, whereas re-orientation of the C^α–H^α bond necessarily affects the conformation of the attached (sometimes bulky) side-chain, which, in turn, may be involved in interactions that are crucial for the overall molecular fold. Futhermore, the C^α–H^α S^2_{MD} values are generally more uniform than the corresponding N-H values, due to the absence of hydrogen bonding propensity.

REMARKS

I. Examiner Interview

On April 11, 2006, Applicant's representative, Stephanie Amoroso, conducted a telephonic interview in connection with the subject application with Examiner Lori A. Clow. According to Applicant's file, during the interview, Applicant's representative and the Examiner discussed the prior art of record and possible claim amendments to clarify the differences between the prior art of record and claim 1. No agreement on allowable subject matter was reached.

II. Claim Status

Claims 2-5 have been cancelled without prejudice or disclaimer subsequent to being withdrawn from consideration by the Examiner as being directed to non-elected subject matter. Applicant reserves the right to pursue the cancelled claims in one or more divisional applications.

Claims 1 and 26, 27, 29 and 30 have been amended to be directed to methods that call for "utilizing the calculated dimensions, chemical and/or electrostatic properties obtained in step b to identify compounds that contain functional groups that can be accommodated [by a cavity on a target protein]." Claim 28 has been similarly amended to be directed a method that calls for "utilizing the calculated dimensions of step b or the chemical and/or electrostatic properties obtained in step c to identify compounds that contain functional groups that can be accommodated [by a cavity on a target protein]." Upon reading the specification, it is self evident that the calculated physical properties of a cavity are used to identify compounds that include functional groups that may be accommodated in the cavity. To the extent that further support is required for "utilizing" such information to identify compounds that bind within an cavity that has been identified on a target protein, such support is found at, e.g., page 12, line 21 et seq., page 18, line 3, et seq., page 25, lines 11-14, and page 27, lines 24-25.

The further amendments to claim 1 are for the purpose of clarity only and do not change the scope or meaning of claim 1. Thus, these further amendments to claim 1 do not add new matter to the application.

By this Amendment, no new matter is added to the application.

Upon entry of this Amendment, claims 1 and 6- 31 are pending.

III. Response to Rejections

The rejections set forth in the Final Office Action are summarized and addressed as follows.

(i) Rejections Under 35 U.S.C. §112, first paragraph (written description). The Examiner has rejected claims 25 and 31 as allegedly including new matter. The rejections are traversed as follows.

With respect to claim 25, the Examiner's position is that the specification fails to provide adequate written description for identifying an allosteric cavity of a target protein using nuclear magnetic resonance, crystal structure analysis, calorimetric values from thermodynamic studies, or computer modeling. The Examiner's position is believed to be mistaken. The specification at page 11, lines 3-5 provides explicit support for using crystal structure analysis, nuclear magnetic resonance, and computer modeling to identify a cavity. Thus, inclusion of these methods in claim 25 does not constitute new matter. With regard to identifying a cavity using calorimetric values from thermodynamic studies, the specification at page 11, lines 3-4 sets forth that the "cavity of the target protein may be identified by any of several well known techniques" (emphasis added) and states explicitly that the methods recited with reference to identifying a cavity are not limiting. The specification at page 10, lines 13-21 discloses that the functionally critical site of a target protein may be inferred from microcalorimetric analysis. Thus, the specification identifies microcalorimetric analysis as a means to analyze protein structure. The disclosure that "any" well known technique may be used to identify a cavity on a target protein, coupled with the explicit disclosure of that microcalorimetric analysis may be used to determine a particular site, i.e., a functionally critical site, is sufficient to show that the inventors had possession of using calorimetric analysis to identify a cavity. Thus, the specification provides adequate written description for the inclusion of "calorimetric values from thermodynamic studies" in claim 25. Ac-

cordingly, the specification includes adequate written description for all of the methods of identifying a cavity that are recited in claim 25. The present rejection should be withdrawn.

With regard to claim 31, the Examiner's position is that the specification fails to provide written description for identifying an allosteric cavity using a technique that includes identifying thermal β-factors. In first response, Applicant notes again that the specification discloses that "any" well known method may be used to identify a cavity (specification at page 11, lines 2-3) and note further that a means of identifying a functionally critical site of a target protein include, "identification of β-factors on the target protein structure as imaged using crystal or nuclear magnetic resonance (NMR) images, either by thermal β-factors on the atoms of target protein from crystal structure or flexible loops inferred from NMR signals" (*see* specification at page 10, lines 13-18). *See also* specification at page 10, lines 20-25 for further discussion of β-factors. When taken together, these passages show possession of using identification of using NMR or crystal structure analysis and further comprising identifying thermal β-factors to identify a cavity. Thus, for at least this reason, the present rejection should be withdrawn.

Moreover, with further regard to "β-factors," one of ordinary skill in the art would understand that identification of β-factors are a basic aspect of crystallography that is part of every crystal study and that β-factors may be calculated from NMR studies. The Examiner is directed, for example, to an enclosed publication, Philippopoulos et al., J. Mol. Biol., 1995 254:771-792, which describes the derivation of β-factors by crystallography and NMR. Thus, one of ordinary skill in the art would understand that the disclosure that a cavity on a target protein may be identified by "crystal structure analysis [or] NMR" (see page 11, line 4) is sufficient to show possession of identifying an allosteric cavity "using nuclear magnetic resonance or crystal structure analysis, and further compris[ing] identifying thermal β-factors," as called for in claim 31. For this reason additionally, the specification provides adequate written description for claim 31.

For at least the reasons set forth above, the specification provides adequate written description for the full scope of claims 25 and 31. Reconsideration of claims 25 and 31 and the withdrawal of the rejection thereof for containing new matter is requested, accordingly.

Table 3. Backbone dynamic parameters

Residue	Secondary structure	S^2_{NMR}	S^2_{MD}	τ_{int} (NMR) (ps)	τ_{int} (MD) (ps)
Leu2		0.51 (0.03)	?0.31 (0.08)	55 (24)	30 (5)
Lys3		0.83 (0.02)	?0.45 (0.06)	1070 (390)	27 (8)
Gln4	β1		?0.59 (0.04)		18 (14)
Val5	β1	0.80 (0.03)	0.89	89 (9)	
Glu6	β1	0.91 (0.02)	0.85 (0.01)		11 (8)
Ile7	β1		0.88 (0.01)		26 (7)
Phe8	β1	0.85 (0.02)	0.90		
Thr9	β1	0.83 (0.03)	0.89 (0.01)		
Asp10	β1	0.72 (0.13)	0.88		
Gly11	β1	0.91 (0.04)	0.90		
Ser12	β1		0.90		
Cys13	β1	0.84 (0.04)	0.79 (0.01)		
Leu14		0.80 (0.02)	0.88	62 (28)	
Gly15		0.62 (0.01)	0.72 (0.03)	1067 (79)	
Asn16		0.74 (0.01)	0.77 (0.01)	4180 (1830)	
Pro17					
Gly18	β2	0.87 (0.01)	0.89 (0.01)		12 (7)
Pro19	β2				
Gly20	β2	0.87 (0.02)	0.90		
Gly21	β2	0.87 (0.02)	0.87		
Tyr22	β2	0.92 (0.02)	0.90		
Gly23	β2	0.86 (0.04)	0.92		
Ala24	β2	0.87 (0.01)	0.89		
Ile25	β2	0.79 (0.03)	0.87		
Leu26	β2	0.87 (0.02)	0.90		
Arg27	β2	0.81 (0.02)	0.90		
Tyr28		0.87 (0.02)	?0.78 (0.02)		35 (10)
Arg29		0.87 (0.02)	?0.31 (0.05)		52 (7)
Gly30		0.83 (0.01)	?0.48 (0.04)	32 (21)	15 (6)
Arg31		0.86 (0.02)	?0.55 (0.10)	1400 (830)	
Glu32	β3	0.83 (0.02)	0.76 (0.02)		15 (12)
Lys33	β3	0.84 (0.01)	0.85 (0.01)		17 (8)
Thr34	β3	0.84 (0.01)	0.81 (0.01)		4 (4)
Phe35	β3	0.83 (0.02)	0.87 (0.01)		7 (5)
Ser36	β3	0.83 (0.01)	0.86		
Ala37	β3	0.82 (0.02)	0.82 (0.01)		
Gly38	β3	0.87 (0.03)	0.87 (0.01)		
Tyr39	β3	0.85 (0.02)	0.91		
Thr40	β3		0.91		
Arg41	β3	0.85 (0.01)	0.89		
Thr42	β3	0.90 (0.01)	0.87 (0.01)		16 (9)
Thr43	αA	0.82 (0.02)	0.89		
Asn44	αA	0.80 (0.05)	0.90 (0.01)		
Asn45	αA		0.91		
Arg46	αA	0.87 (0.02)	0.92		
Met47	αA	0.87 (0.02)	0.91		
Glu48	αA		0.93		
Leu49	αA		0.91		
Met50	αA	0.91 (0.04)	0.93		
Ala51	αA		0.92		
Ala52	αA	0.82 (0.02)	0.90		
Ile53	αA		0.87		
Val54	αA		0.92		
Ala55	αA	0.86 (0.02)	0.89		
Leu56	αA	0.79 (0.05)	0.89		
Glu57	αA		0.91 (0.01)		
Ala58	αA	0.87 (0.01)	0.88 (0.01)		22 (11)
Leu59		0.78 (0.02)	0.82 (0.02)		
Lys60			?0.63 (0.11)		51 (40)
Glu61		0.74 (0.01)	0.71 (0.06)	28 (11)	47 (33)
His62		0.82 (0.01)	0.88 (0.01)		
Cys63			0.86 (0.01)		
Glu64	β4		0.88		4 (2)
Val65	β4	0.76 (0.03)	0.88 (0.01)		4 (2)
Ile66	β4	0.86 (0.03)	0.90		
Leu67	β4	0.88 (0.03)	0.91		
Ser68	β4	0.88 (0.03)	0.90		
Thr69	β4	0.93 (0.04)	0.89		4 (2)
Asp70			0.90		3 (1)
Ser71	αB	0.60 (0.08)	0.88 (0.01)		
Gln72	αB	0.78 (0.06)	0.86 (0.01)		4 (2)

continued

Table 3. *continued*

Residue	Secondary structure	S^2_{NMR}	S^2_{MD}	τ_{int} (NMR) (ps)	τ_{int} (MD) (ps)
Tyr73	αB		0.90		
Val74	αB		0.92		
Arg75	αB	0.95 (0.02)	0.90		
Gln76	αB		0.92		
Gly77	αB		0.91		
Ile78	αB	0.86 (0.02)	0.92		
Thr79	αB		0.90		
Gln80	αB		0.86 (0.01)		2 (1)
Trp81	αC	0.83 (0.02)	0.85 (0.01)		2 (1)
Ile82	αC	0.84 (0.03)	0.92		5 (2)
His83	αC	0.90 (0.02)	0.91		
Asn84	αC	0.85 (0.01)	0.86 (0.01)		
Trp85	αC		0.90		4 (2)
Lys86	αC	0.88 (0.02)	0.91		5 (2)
Lys87	αC	0.87 (0.01)	0.86 (0.01)		5 (3)
Arg88	αC		0.89		
Gly89		0.80 (0.02)	0.82 (0.01)		2 (1)
Trp90		0.86 (0.02)	0.90		
Lys91		0.79 (0.01)	0.82 (0.05)	37 (14)	
Thr92		0.82 (0.01)	0.86 (0.02)	31 (17)	
Ala93		0.83 (0.03)	?0.78 (0.03)		24 (9)
Asp94		0.77 (0.01)	?0.79 (0.02)	48 (10)	19 (6)
Lys95		0.83 (0.01)	0.80 (0.02)	36 (17)	
Lys96		0.80 (0.01)	0.78 (0.02)	1110 (250)	15 (5)
Pro97					
Val98		0.79 (0.01)	0.82 (0.01)	45 (20)	4 (2)
Lys99		0.84 (0.01)	0.85 (0.01)		10 (3)
Asn100	αD	0.85 (0.02)	0.87 (0.01)		8 (4)
Val101	αD	0.81 (0.01)	0.89		4 (3)
Asp102	αD	0.85 (0.01)	0.90		
Leu103	αD	0.80 (0.02)	0.87		3 (1)
Trp104	αD		0.91		
Gln105	αD	0.81 (0.02)	0.92		
Arg106	αD	0.85 (0.02)	0.90		
Leu107	αD	0.82 (0.04)	0.92		
Asp108	αD		0.92		
Ala109	αD	0.82 (0.02)	0.90		
Ala110	αD	0.82 (0.02)	0.86		2 (1)
Leu111	αD	0.81 (0.02)	0.90		
Gly112	αD	0.86 (0.02)	0.77 (0.02)		
Gln113			0.71 (0.03)		12 (3)
His114		0.79 (0.03)	0.81 (0.01)		5 (3)
Gln115	β5		0.86 (0.01)		
Ile116	β5	0.74 (0.01)	0.82 (0.01)		
Lys117	β5	0.84 (0.02)	0.89		
Trp118	β5	0.73 (0.02)	0.85 (0.01)		2 (1)
Glu119	β5	0.84 (0.03)	0.88		2 (1)
Trp120	β5	0.74 (0.06)	0.87		
Val121		0.93 (0.06)	0.84 (0.01)		
Lys122		0.86 (0.05)	0.84 (0.02)		
Gly123		0.54 (0.07)	?0.57 (0.12)	162 (128)	
His124			?0.68 (0.02)		25 (13)
Ala125		0.77 (0.03)	0.80 (0.03)	84 (36)	
Gly126		0.74 (0.09)	?0.24 (0.05)		24 (18)
His127	αE	0.86 (0.09)	0.69 (0.03)		
Pro128	αE				
Glu129	αE	0.90 (0.05)	0.91		
Asn130	αE	0.83 (0.05)	0.89		
Glu131	αE		0.89		
Arg132	αE	0.83 (0.01)	0.90		3 (2)
Cys133	αE	0.91 (0.03)	0.92		5 (4)
Asp134	αE	1.00 (0.01)	0.92		
Glu135	αE		0.90		
Leu136	αE	0.89 (0.01)	0.91		
Ala137	αE		0.92		
Arg138	αE		0.93		
Ala139	αE	0.89 (0.02)	0.89		9 (8)
Ala140	αE	0.88 (0.01)	0.90		
Ala141	αE	0.92 (0.03)	0.89		
Met142	αE	0.78 (0.02)	0.81 (0.01)		22 (11)

continued

Table 3. *continued*

Residue	Secondary structure	S^2_{NMR}	S^2_{MD}	τ_{int} (NMR) (ps)	τ_{int} (MD) (ps)
Asn143		0.83 (0.01)	?0.68 (0.03)		19 (11)
Pro144					
Thr145		0.80 (0.02)	?0.82 (0.03)		13 (8)
Leu146		0.88 (0.02)	0.79 (0.03)		
Glu147		0.80 (0.01)	0.85 (0.01)		
Asp148		0.81 (0.01)	0.88		4 (1)
Thr149			?0.75 (0.01)		6 (2)
Gly150		0.79 (0.01)	0.84 (0.01)		6 (3)
Tyr151		0.85 (0.01)	0.75 (0.06)		
Gln152		0.80 (0.01)	0.72 (0.03)	1470 (190)	10 (5)
Val153		0.77 (0.01)	0.77 (0.03)	1200 (51)	14 (7)
Gln154		0.78 (0.01)	0.75 (0.02)	1070 (24)	13 (4)
Val155		0.65 (0.01)	?0.21 (0.08)	755 (12)	25 (5)

S^2_{NMR} in the order parameter determined by NMR (Mandel *et al.*, 1995; in cases for which both S^2 and S^2_f have been determined experimentally, only S^2_f, the order parameter of motions with correlation times $\tau_{int} < 50$ ps, is shown); S^2_{MD} is the order parameter computed from the simulation. A question mark before the value indicates that the N–H correlation function did not reach a well-defined plateau value within 200 ps correlated time. τ_{int} (NMR) is the effective correlation time of internal motion determined by NMR (in some cases it reflects motion on a timescale that is longer than the simulation length, e.g. Arg31). τ_{int} (MD) is the correlation time of internal motion computed from the simulation. Uncertainties of all quantities are shown in parentheses, theoretical uncertainties are determined by the jackknife method described in Methods, and values < 0.005 are not shown. Only τ_{int} (MD) values higher than 1 ps are shown. τ_{int} (MD) values whose theoretical uncertainty is equal to or higher than their absolute value (due to extreme noise in the auto-correlation function) are not shown.

Effective correlation times

Calculated effective correlation times, τ_{int}, based on the 400 ps production trajectory are listed in Table 3. Only τ_{int} values greater than 1 ps are shown, due to the coordinate sampling interval of 100 fs (see Methods) that has resulted in re-orientational correlation functions with low precision on the fs timescale. The τ_{int} values were computed from equation (7), with T defined as the point at which the correlation function matched the corresponding S^2_{MD} value (calculated from equation (6)) to within 0.001. For the majority of residues that have been experimentally assigned a motional timescale (Mandel *et al.*, 1995), τ_{int} values seldom exceed 20 ps (the sensitivity lower limit of the NMR experiment). For this reason, the choice of 40 ps for the block size used in the "jackknife" algorithm (see Methods) for estimation of the statistical uncertainties of S^2 and τ_{int} values from the MD trajectory is justified in most cases.

A relatively small number of τ_{int} values were determined experimentally, due to the conservative nature of the fitting protocol and low sample concentrations that gave rise to nuclear Overhauser enhancement (NOE) values of low precision (Mandel *et al.*, 1995). This fact, along with the limited precision to which τ_{int} values can be computed from the 400 ps trajectory, makes comparison between theoretical and experimental τ_{int} values problematic. However, there are some interesting patterns. Most αC residues (positions 83 to 88) are assigned a τ_{int}, in contrast to neighbouring αB (Table 3). This behaviour is consistent with the relatively high B-factors of residues 83 to 88 (Table 1). Most residues of termini, loops and turns possess a τ_{int}, which is computed to be lower than 25 ps in most cases. Finally, long stretches of residues in αA, αB and αD (positions 104 to 109) are not assigned a motional timescale τ_{int} from the present simulation, indicating severely restricted librational motion of these N–H bonds with $\tau_{int} < 1$ ps.

Discussion

Comparison with X-ray data

X-ray isotropic B-factors reflect protein mobility on two distinct timescales: 10^{-15} to 10^{-11} seconds (fast thermal fluctuations) and 10^{-9} to 10^{3} seconds (conformational changes). Since only highly populated conformers are observable *via* X-ray diffraction, the method is relatively insensitive to motions, such as domain hinge-bending or aromatic ring flips, that occur either too slowly or too infrequently to yield distinct conformers and hence detectable electron densities (Petsko & Ringe, 1984). Currently, only motions on the 10^{-15} to 10^{-10} seconds timescale are observable in common by X-ray diffraction and molecular dynamics simulations; this type of fluctuation is expected to dominate the backbone dynamics of regular secondary structure elements that are part of tightly packed protein hydrophobic cores. In general, the observed mean-square atomic displacement, $\langle x^2 \rangle$, can be written as a sum of three contributions (Petsko & Ringe, 1984);

$$\langle x^2 \rangle = \langle x^2 \rangle_v + \langle x^2 \rangle_c + \langle x^2 \rangle_l \quad (1)$$

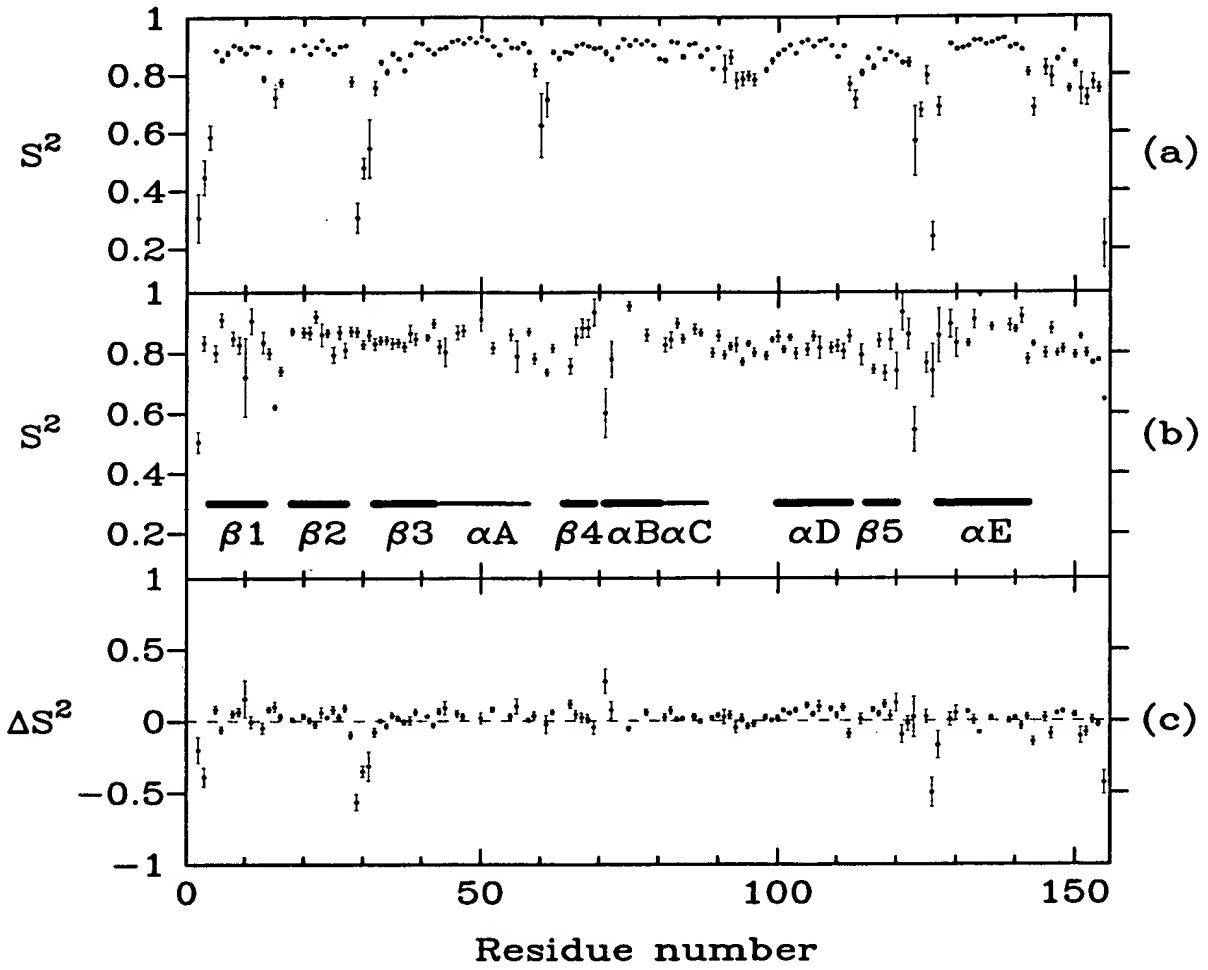

Figure 4. (a) Generalized order parameters S^2 of the backbone N–H bonds calculated from the 100 to 500 ps portion of the molecular dynamics trajectory. (b) Corresponding NMR data. (c) Difference between calculated and experimental values.

where $\langle x^2 \rangle_v$ is the contribution due to thermal fluctuations (vibrations within a single harmonic potential well), $\langle x^2 \rangle_c$ is the contribution due to static and dynamic conformational disorder (molecules existing in different conformations in the crystal) and $\langle x^2 \rangle_l$ is the contribution due to lattice disorder (crystal heterogeneity).

To make a quantitative comparison of the X-ray *B*-factors with the computed values, an attempt to subtract out the lattice disorder contribution was made by translating the experimental values so that the lowest value was equal to the corresponding theoretical value. Thus, the 2RN2 data were shifted by 3.71 Å^2 downward, whereas the 1RNH data were not translated, as the lowest 1RNH value was lower than the corresponding computed one. The comparative profiles of the calculated and experimental main-chain *B*-factors averaged on a per residue basis in Figure 6 show that the overall qualitative features of the experimental data are reproduced well by the simulation. The largest discrepancy between theory and experiment is found in the reverse turn of the handle, where the *B*-factors of residues 92 to 95, especially Ala93, are much higher than the corresponding X-ray values. On the other hand, there is good agreement between S^2_{NMR} and S^2_{MD} values for these residues (Table 3). Since Lys91, Ala93 and Asp94 are involved in a number of intermolecular hydrogen bonding contacts in the 2RN2 structure, the removal of these constraints during molecular dynamics may have resulted in higher atomic fluctuations for the reverse turn relative to those observed in the X-ray structure. In fact, this region underwent a significant displacement (by about 5 Å with respect to the 2RN2 coordinates) in a direction away from the protein core during equilibration. As a result, this region became exposed to vacuum, thereby undergoing large atomic fluctuations relative to the rest of the protein. Therefore, the above crystal contacts may have had an important effect on the structure and dynamics of the reverse turn.

Apart from the handle region and the termini, the computed average *B*-factors are generally lower than the X-ray values. Some possible reasons for the observed systematic error are: (1) the least-squares fitting procedure employed in separating internal motions from overall molecular diffusion (inadver-

tently removing some internal motion). (2) The limited timescale of the simulation (limiting exploration of conformational space, thus decreasing the computed values). (3) Inadequacies in the empirical force-field (leading to an underestimation of mobility). (4) Under-correction of the lattice disorder in the crystal (increasing the experimental values). (5) Under-correction of the reflections for radiation damage (causing systematic overestimation of *B*-factors; Petsko & Ringe, 1984). (6) Phonon vibrations in the crystal that would be absent in a solution environment (resulting in enhanced X-ray *B*-factors).

Comparison of *B*-factors from the simulation and the 1RNH structure reveals close agreement for several residues in αA, αB and αC, which also contain some residues with X-ray *B*-factors that are lower than the computed values. Thus, the least-squares fitting procedure employed in (1), above, does not seem to be a significant contribution to the observed discrepancy. Comparison of the computed *B*-factors with those from a 200 ps production trajectory yields significant differences only for some of the turns (such as residue 60), whose *B*-factors are increased upon doubling the trajectory; this, in conjunction with the NMR finding of internal correlation times only in irregular secondary structure elements (Mandel *et al.*, 1995), suggests that the limited timescale of the simulation as in (2), above, contributes to at least part of the observed difference between computed and X-ray *B*-factors for the loops, turns and termini.

However, it seems unlikely that the significant differences between experimental and computed *B*-factors of regular secondary structure elements stem from conformational disorder in the crystal structure not sampled by the present trajectory. Such large-scale motions would probably be inhibited, especially in the tightly packed hydrophobic core of strands β1 to 4. Comparison of the 1RNH and 2RN2 values (Figure 6 and Table 1) suggests that the different experimental procedures (multiwavelength anomalous diffraction *versus* multiple isomorphous replacement, respectively) and refinement protocols (Yang *et al.*, 1990; Katayanagi *et al.*, 1990) may contribute to the observed discrepancy. Furthermore, for a few proteins like ribonuclease A (Petsko, 1975), crambin (Teeter & Hendrickson, 1979), and avian pancreatic polypeptide (Glover *et al.*, 1983), which diffract to resolutions comparable to that of small molecule crystals, i.e. 1 Å or higher, a number of atoms have *B*-factors as low as 3-4 Å^2, similar to those observed for atoms in small organic crystal structures. Other small, well-diffracting proteins like rubredoxin (solved to 1.2 Å resolution) have *B*-factors as low as 7 Å^2 (Petsko & Ringe, 1984), comparable to the average computed *B*-factors of the β strands (6.3 Å^2) and α helices (7.3 Å^2) of RNase H$_I$. Taken together, the large differences between the 1RNH and 2RN2 values, as well as the relatively high 2RN2 *B*-factors for the regular secondary structure elements, indicate that the observed discrepancy in the β strands and α

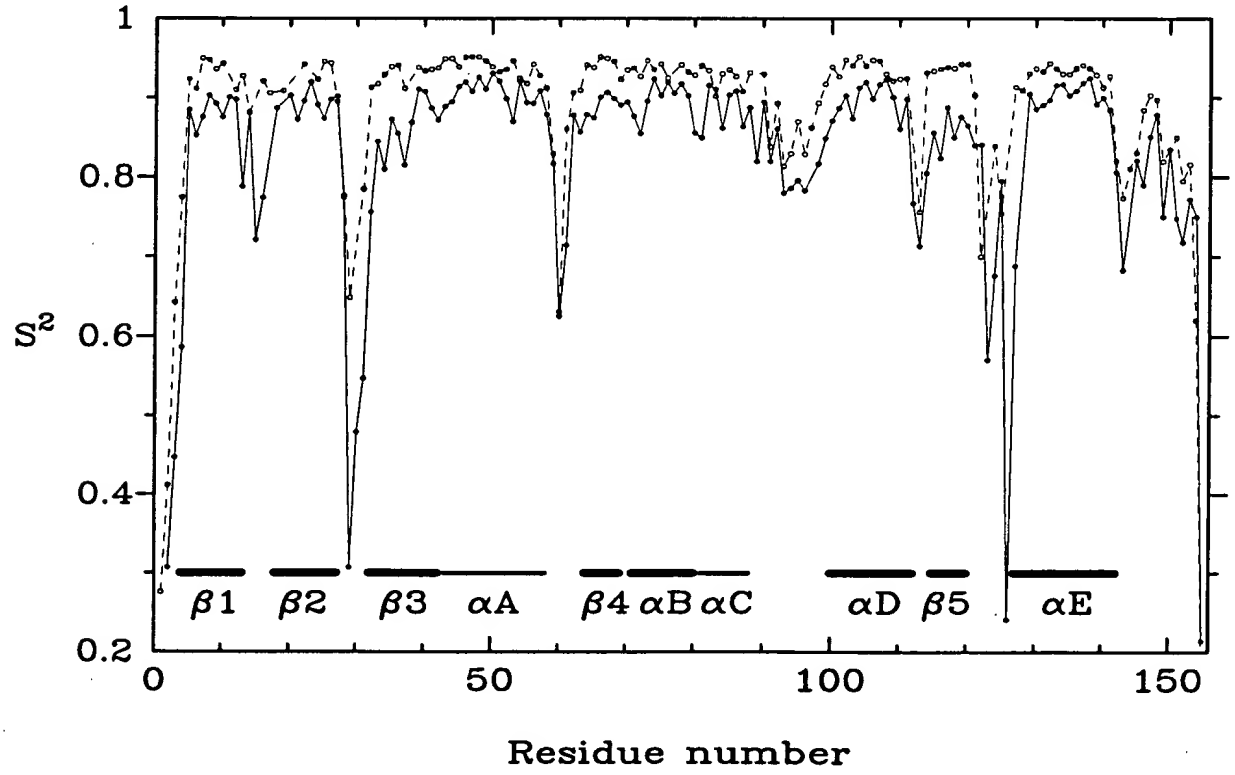

Figure 5. Comparison of the generalized order parameters S^2 of the 141 backone C^α–H (broken line) *versus* N–H bonds (continuous line) calculated from the 100 to 500 ps portion of the molecular dynamics trajectory.

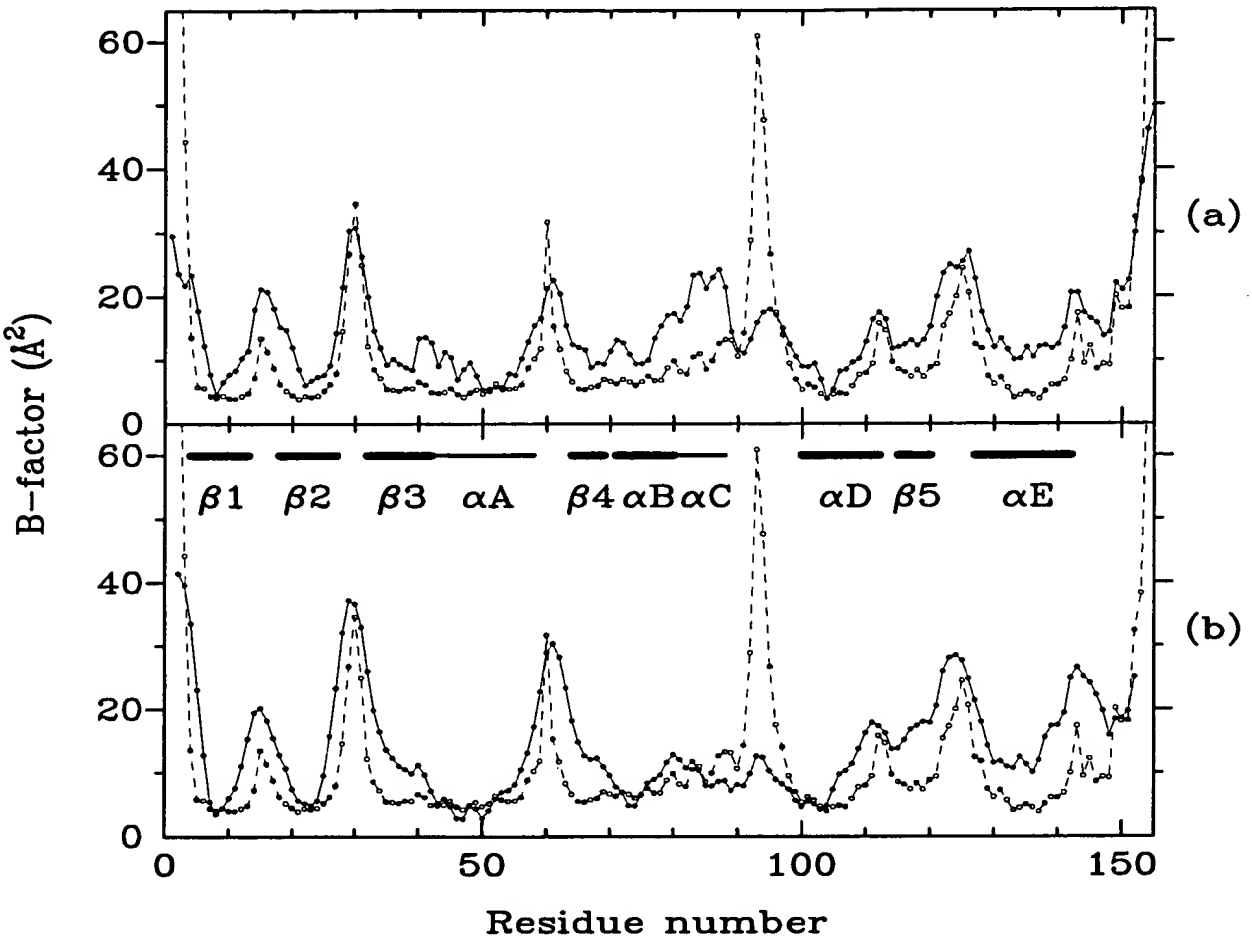

Figure 6. (a) Comparison of calculated main-chain *B*-factors averaged on a per residue basis (broken line) with experimental values (continuous line) taken from PDB entries 2RN2. (b) With 1RNH. The 2RN2 values in (a) are downshifted by 3.71 Å^2 (see the text).

helices may result from factors (4) to (6) stated above.

Comparison with NMR data

In analogy to the computed *B*-factors, the S^2_{MD} values are generally overestimated relative to the S^2_{NMR} values of most residues (see Figure 4). The major cause of the observed "systematic" error does not stem from the least-squares fitting procedure employed in removing tumbling, since correlation functions in which overall tumbling was not removed were very similar to those obtained after the least-squares fitting (plateau values were within 0.02); this indicates that on the simulation time-scale, the computed order parameters are insensitive to the manner in which overall translational and rotational diffusion is treated. Another possible reason for the discrepancy may stem from the assumption of a fixed N–H bond length of 1.02 Å inherent in the experimental fitting procedure by which NMR relaxation data were fitted to spectral density models containing one, two or three model-free dynamical parameters (Mandel *et al.*, 1995). In contrast, bond oscillations have been allowed to contribute to the S^2_{MD} values presented in Table 3, as the SHAKE algorithm (van Gunsteren & Berendsen, 1977) that permits longer simulation time steps by fixing covalent bonds to hydrogen atoms was not used in the present work. Comparison between S^2_{MD} values shown in Table 3 and values computed with a fixed N–H distance of 1.02 Å shows that the latter are higher by about 0.005. This is, therefore, not a significant contribution to the observed differences. A more significant contribution may be due to the neglect of zero-point vibrational motion of the N–H and C$^\alpha$–H bond. This is suggested by backbone heteronuclear S^2 values for BPTI (Bruschweiler, 1992) and a zinc-finger peptide (Palmer & Case, 1992). A systematic difference between classical statistics (as in this work) and quantum statistics was found to be between 5% and 6%, independent of the amino acid type and position in the protein; i.e. order parameters from a classical analysis were greater than those from a quantum mechanical analysis by an average of 0.05. In the following, possible reasons for significant discrepancies between MD- and NMR-derived order parameters (greater than 0.05), as well as the physical basis of unusual or

interesting patterns of the observed dynamics, are discussed.

α-Helices and β-strands

Table 4 lists average S^2 values per secondary structure element and compares them with values determined by NMR (Mandel *et al.*, 1995). Average S^2_{MD} values of residues in β strands and α helices typically fluctuate around 0.89(±0.02). There is reasonably good agreement between the theoretical and experimental average S^2 values of the α helices and β strands: the difference between the two sets of mean S^2 values is less than 0.05 (the estimated correction for zero-point vibrational motion) for four of the five β strands (β1 to 4) and three of the five α helices (αB, αC, and αE). It is somewhat higher (0.06–0.07) for αA and αD, and is as high as 0.10 for β5. In all cases, the mean S^2 values for the regular secondary elements are overestimated relative to experiment.

β5 strand

The average S^2_{NMR} of β5 (0.77(±0.01)) is unusual, although β-sheet S^2 values in the neighbourhood of 0.80 have been reported for other proteins (Clore *et al.*, 1990; Stone *et al.*, 1992; Farrow *et al.*, 1994). The NMR relaxation study (Mandel *et al.*, 1995) found that β5 residues (116 to 120) exhibit a pattern of alternate high and low mobility with an average S^2 of 0.74(±0.01) for Ile116, Trp118 and Trp120 and 0.84(±0.02) for Lys117 and Glu119. The simulation has reproduced this alternating motional pattern, and the S^2_{MD} values of Lys117 and Glu119 are close to the corresponding experimental values. The differential mobility of the β5 residues correlates with the hydrogen bonding patterns of the backbone nitrogen atoms. The N–H mobility of Gln115, Lys117 and Glu119 is restricted, since their amide hydrogens interact with the carbonyl oxygens of Cys63, Val65 and Leu67, respectively. In contrast, there are no stable protein or water hydrogen bonds to the backbone amide hydrogens of Ile116, Trp118 and Trp120, allowing their N–H bonds to exhibit greater mobility. The only analogous structural situation to residues 116 to 120 occurs at the opposite edge of the β-sheet (strand β3), where the amide hydrogens of Glu32, Thr34 and Ser36 are not involved in hydrogen bonds, whereas those of Lys33 and Phe35 interact with the carbonyl oxygens of Leu26 and Ala24 of β2, respectively. However, an alternating dynamical pattern for these residues was not observed *via* NMR relaxation: the S^2_{NMR} of residues 32 to 36 are uniform (0.83(±0.02)). The simulation, in contrast, does indicate a pattern of alternate high mobility for Glu32, Thr34 and Ser36 and lower mobility for Lys33 and Phe35 (see Table 3): in particular, the S^2_{MD} differences between Lys33 and Glu32 (0.09) and between Phe35 and Thr34 (0.06) are greater than the uncertainties.

Table 4. Weighted mean order parameters for secondary structure elements

Secondary structure	Residue range	S^2_{NMR}	S^2_{MD}	Residues quantified
β1	4–13	0.85 (0.01)	0.88	7
β2	18–27	0.87 (0.01)	0.89	9
β3	32–42	0.85 (0.01)	0.87	10
β4	64–69	0.85 (0.01)	0.90	5
β5	115–120	0.77 (0.01)	0.87	5
⟨β⟩		0.84	0.88	
αA	43–58	0.86 (0.01)	0.91	9
αB	71–80	0.91 (0.01)	0.91	4
αC	81–88	0.87 (0.01)	0.90	6
αD	100–112	0.83 (0.01)	0.90	11
αE	127–142	0.89 (0.01)	0.90	11
⟨α⟩		0.87	0.91	
$Turn_{12}$	14–17	0.70 (0.01)	0.85	3
$Turn_{23}$	28–31	0.86 (0.01)	0.66 (0.02)	4
$Loop_{A4}$	59–63	0.77 (0.01)	0.86 (0.01)	3
$Handle_{CD}$	89–99	0.81	0.85	10
$Loop_{5E}$	121–126	0.80 (0.02)	0.82 (0.01)	5
⟨*loops*⟩		0.79	0.84	
C Terminus	143–155	0.78	0.82	11

The last column shows the number of residues in that element quantified experimentally. In some instances, S^2_{NMR} values exist for a limited number of residues, e.g. in αB. S^2_{MD} values and their uncertainties were computed as weighted means, including only S^2 of residues that have been quantified experimentally. S^2_{NMR} for α helices and β strands are taken from Mandel *et al.* (1995), Table 4, whereas values for non-regular secondary structure elements are computed as weighted averages of S^2 from Table 1 of the same reference. In cases where both an S^2 and an S^2_f value are reported experimentally, the S^2_f value has been chosen for the averaging. ⟨β⟩, ⟨α⟩ and ⟨*loops*⟩ are average S^2 values over the five β strands, five α helices and five loops, respectively.

Coiled-coil interface

Significant discrepancies between the NMR data and the simulation are seen for several residues in the coiled-coil interface between helices αA and αD (Yang *et al.*, 1990; Katayanagi *et al.*, 1992). The interactions between Arg46 of αA and Asn100 and Asp102 of αD, as well as between Glu57 of αA and Arg106 of αD observed in the 2RN2 structure, are all preserved in the simulation. NMR-derived N–H order parameters for this region indicate a pattern of reduced mobility for the charged residues Arg46, Asp102 and Arg106 (mean $S^2_{NMR} = 0.85(\pm 0.01)$) and of enhanced mobility for the hydrophobic residues, Leu56, Leu103, Leu107 and Leu111 (mean $S^2_{NMR} = 0.80(\pm 0.01)$); no data have been reported for the other residues involved in stabilizing the interface, Leu49, Ile53 and Glu57. In contrast, the MD-derived backbone N–H order parameters for the charged residues, Arg46, Glu57, Asp102 and Arg106 (mean $S^2_{MD} = 0.91(\pm 0.01)$) are similar to those for residues involved in hydrophobic interactions between αA and αD: Leu49, 56, 103, 107, 111 and Ile53 (mean $S^2_{MD} = 0.89(\pm 0.01)$). Relative to the experimental values, the S^2_{MD} values are overestimated by about 0.05 for the charged residues and roughly 0.1 for the hydrophobic ones. Although order parameters are only sensitive to ps-ns type motions, individual values may still be ensemble averages over several conformational substrates inter-converting on timescales longer than the ps-ns

timescale. As significant conformational exchange terms have been identified for most of the residues, especially all the leucines comprising αA and αD (Mandel *et al.*, 1995), the lower S^2 observed experimentally may reflect the existence of additional less rigid conformations in equilibrium with the one sampled in the trajectory and interconverting on the μs-ms timescale. In particular, there may be a coiled-coil conformational state where the hydrophobic residues stabilizing the interface may not be well packed (thus giving them an extra degree of mobility), but the overall structure of αA and αD is maintained (thus not affecting the mobility of the charged residues).

αC helix

The S^2_{NMR} and S^2_{MD} values for αC are in close agreement (Table 3). Both theory and experiment indicate that the mean αC S^2 values (Table 4) do not deviate from the average S^2 values expected for regular secondary structure, whereas the mean αC *B*-factors (Table 1) are the highest among all regular secondary structure elements (although the *B*-factors in 1RNH do not indicate high backbone mobility for αC). This difference between X-ray *B*-factors and NMR order parameters illustrates the variance in dynamical information that the two sets of data can provide. *B*-factors are directly related to atomic rms fluctuations. N–H order parameters, on the other hand, are only sensitive to angular fluctuations and the local environment of the N–H bond, i.e. proximity to hydrogen-bond partners, and may therefore reflect more localized motions as compared to *B*-factors.

Several protein–protein hydrogen bonds involving αC residues, not present in the X-ray structure, are formed during the simulation: Trp81(O)–Asn84(N), His83(O)–Lys87(N), Arg88(N^{ϵ})–Lys91(O) and Arg88($N^{\eta 1}$)–Trp90(O). The absence of the His83(O)–Lys87(N) interaction from the 2RN2 structure (normally expected to be present as part of the α helix) may have been due to the intermolecular interactions of Trp85, Lys86, Lys87 and Gly89 with Gln115, Gly123 and His124 of an adjacent molecule (Katayanagi *et al.*, 1992). The presence of hydrogen bonding partners for the amide hydrogens of all the αC residues (except for His83) probably results in the relatively high S^2 values of helix C. Interestingly, despite the additional interactions found in the simulation, the computed *B*-factors for the αC residues (especially 83 to 88) are higher than average *B*-factors for the other α helices; this trend is in accord with the 2RN2 *B*-factors. Thus, the observed high backbone *B*-factors of αC may be the result of collective translational displacements of the helix, in which the individual N–H bonds preserve their orientations, thus accounting for the relatively high S^2_{NMR} and S^2_{MD} values.

Loops and the C terminus

The average S^2 of the two isolated reverse turns (residues 14 to 17 and 28 to 31), two loops (residues 59 to 63 and 121 to 126), the handle region (residues 89 to 99) and the C-terminal region (residues 143 to 155) are listed in Table 4 and compared with corresponding experimental values. The average S^2 values of residues in turns, loops and the C terminus are lower than those in the β strands and α helices. This is consistent with the fact that an S_f^2 and/or a motional timescale have been assigned experimentally to residues belonging exclusively to non-regular secondary structure elements, reflecting the enhanced flexibility of these regions. For two of the three loops that have been implicated in binding, the β5–αE loop and the handle region, as well as the C-terminal region, the average theoretical and experimental S^2 values are in reasonable accord (within 0.04). Agreement between theory and experiment for the loop between αA and β4 is better than the average S^2 values indicate: two of the three S^2_{NMR} values (for Leu59 and Glu61) are in accord with the individual S^2_{MD} values (within the uncertainties) and only the S^2_{NMR} of His62, which had the second lowest goodness-of-fit (11.83) in the protein, was found to differ significantly from the computed number (by 0.06). However, the β2–β3 turn, which does not contain any residues implicated in binding and/or catalysis, and the β1–β2 turn, have average S^2_{MD} values that differ significantly from the corresponding S^2_{NMR} values (by 0.14 and 0.20, respectively); the discrepancy is probably due to errors in the simulation (insufficient sampling and possibly inadequate solvation) and/or the experimental fitting protocol (low goodness-of-fit values).

Handle region

Agreement between the S^2_{NMR} and S^2_{MD} values for the handle region is one of the best in the protein, as shown in Tables 3 and 4. Trp90 has the highest S^2_{NMR} and S^2_{MD} value in the handle region. Consistent with the NMR data, the X-ray backbone temperature factors of Trp90 are among the lowest for residues 81 to 99 in the 2RN2 structure (Katayanagi *et al.*, 1992), whereas in the 1RNH structure (Yang *et al.*, 1990) rigidity is less exclusive to Trp90. In the 2RN2 structure, Trp90 interacts with αC and the handle region on the other side of the reverse turn *via* two hydrogen bonds, Trp90(H^N)–Trp85(O) and Trp90(O)–Val98(H^N) (Katayanagi *et al.*, 1992). Only the former interaction is maintained during the simulation, so that the latter is not responsible for the NMR-observed rigidity of the Trp90 backbone (Mandel *et al.*, 1995). In addition, the carbonyl oxygen of Gly89 forms a hydrogen bond with Lys91(H^N), thereby further rigidifying the Trp90 N–H vector through the Gly89–Trp90 peptide bond. The Gly89(O)–Lys91(H^N) hydrogen bond is not seen in the 2RN2 crystal structure (Katayanagi *et al.*, 1992). Note that the Trp90(H^N)–Trp85(O) and

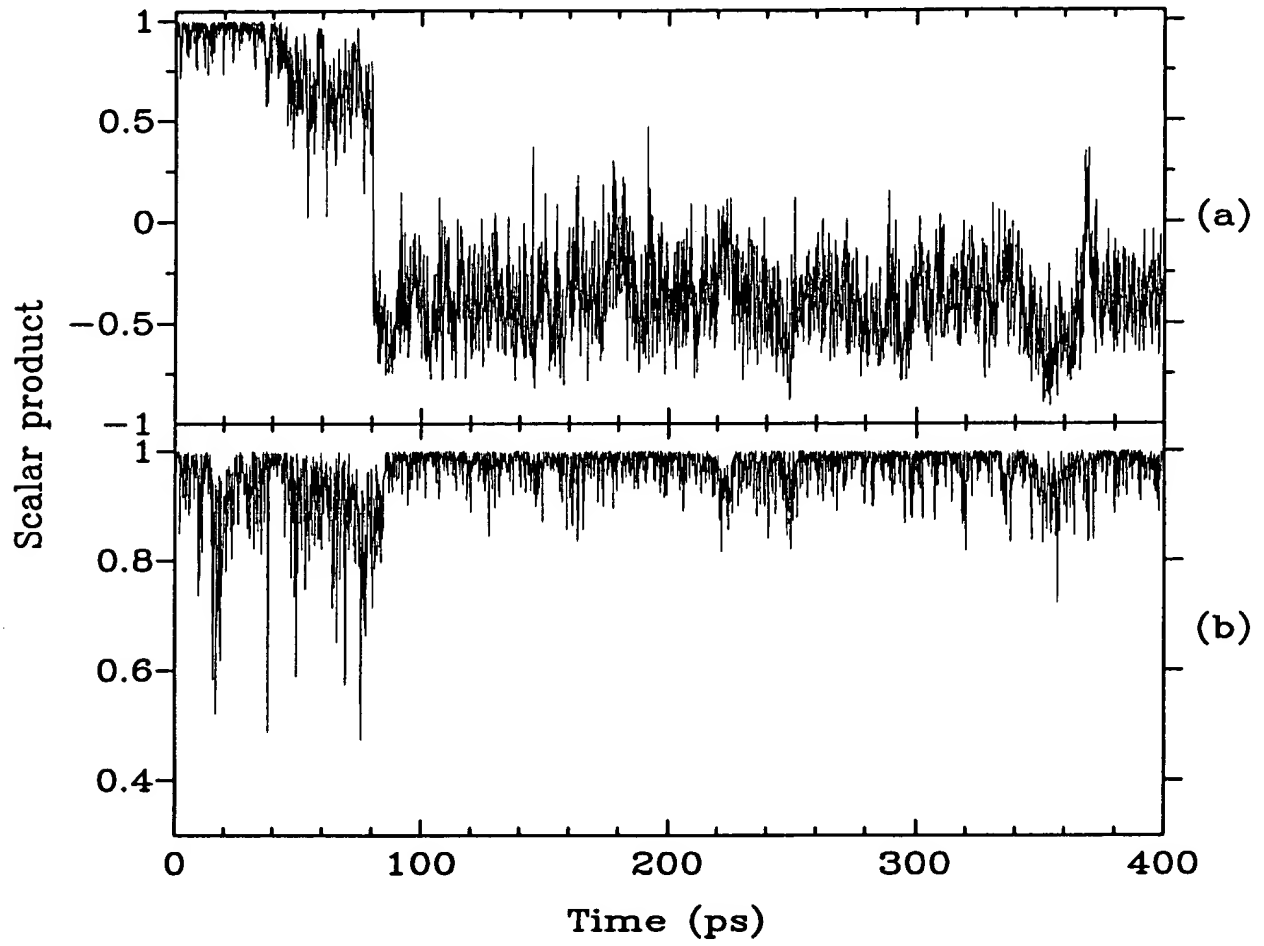

Figure 7. (a) The scalar products of the Gly123 and (b) Ala125 N-H bonds relative to their orientation at $t = 0$ of the production trajectory as a function of simulation time.

Gly89(O)-Lys91(H^N) interactions do not appear to be screened by solvent and furthermore, the backbone and side-chain atoms of Trp90 do not form stable hydrogen bonds, with water molecules.

β5–αE loop

The average S^2_{NMR} and S^2_{MD} values of this loop are within the experimental uncertainty (Table 4); the biggest discrepancy in the two sets of values is due to Gly126, whose N-H auto-correlation function did not decay to a plateau value, causing its S^2_{MD} to differ significantly from the experiment. The origin of the relatively significant τ_{int} values (162(±128) and 82(±36) ps) found experimentally for Gly123 and Ala125 (Mandel *et al.*, 1995), the only residues in the β5-αE loop assigned a motional timescale, may be linked to a concerted transition in the scalar products of the Gly123 and Ala125 N-H bonds relative to their initial orientation at about 80 ps into the production trajectory (Figure 7). This transition coincides with sharp transitions at 80 to 85 ps in the Lys122(O)-Ala125(N) distance as well as the Gly123 ϕ and ψ dihedral angles (Figure 8). At $t = 85$ ps, a water molecule (Bul3397) forms a hydrogen bond with Ala125(H^N). Another water molecule (Bul3586) comes within hydrogen-bonding distance of His124(O) at $t = 90$ ps, and at $t = 112$ ps, the same water molecule also forms a hydrogen bond with Glu131($O^{\epsilon 2}$). Both these interactions remain stable for the remainder of the simulation. The water bridge His124(O)-Bul3586-Glu131($O^{\epsilon 2}$) (and to a lesser extent, Ala125(H^N)-Bul3397) is responsible for restricting the mobility of the Ala125 N-H bond, due to the partially double-bond nature of the peptide bond connecting His124 C = O and Ala125 N-H. Finally, a third water molecule (Bul3315) comes to contribute to the stabilization of the backbone conformation of His24 and Ala125 by forming hydrogen bonds simultaneously with their amide hydrogen atoms at approximately $t = 250$ ps. As a precaution against artifacts due to insufficient equilibration giving rise to the transition at ca 85 ps, the S^2 of Gly123 was recalculated from the last 300 ps of the trajectory and found to be 0.77(±0.03), i.e. much higher than the value of 0.57(±0.12) calculated including the transition and the experimental value of 0.54(±0.07). As there are no exchange terms found in the β5-αE loop, implying that ps-ns motions are probably the only contribution, this transition

appears to account for the S^2_{NMR} values of Gly123 and Ala125, and thus the dynamics of this loop.

Conserved residues

Of the seven most highly conserved residues in RNase H_I, backbone dynamical parameters have been determined experimentally only for Asp10, Ser71, Asp130 and Asp134 (Mandel *et al.*, 1995); the other three catalytically important residues could not be quantified due to exchange line broadening (for Glu48 and Asp70) and resonance overlap (for His124). Only Asn130 was found to have an unexceptional S^2_{NMR} (0.83($\pm$0.05)), and the only one in accord with theory (S^2_{MD} = 0.89($\pm$0.01)).

Acidic residues

The S^2_{NMR} value for Asp10 has the highest uncertainty in the protein (0.72($\pm$0.13)), whereas the S^2_{NMR} value for Asp134 is the highest in the protein (1.00($\pm$0.01)). The S^2_{MD} value for Asp10 (0.88), near the upper limit of the experimental number (0.85), is consistent with the occurrence of this residue as part of the regular secondary structure and its low computed and X-ray *B*-factors; thus, this suggests that the S^2_{NMR} value may be an overestimate of backbone mobility. The S^2_{MD} value of Asp134 (0.92) is significantly lower than the S^2_{NMR} value. In addition, the S^2_{MD} values of neighbouring Cys133 (0.92), Glu135 (0.90) and Leu136 (0.91) are also lower than the S^2_{NMR} value of Asp134. Since the S^2_{MD} values of Cys133 and Leu136 (Glu135 was not quantified experimentally) are close to their experimental values, it is possible that the S^2_{NMR} value of Asp134 may be in error. Glu48 possesses the highest S^2_{MD} value (0.93) among the four absolutely conserved acidic residues whereas Asp70, the single residue between β4 and αB, has a S^2_{MD} (0.90) in between the average S^2_{MD} values of a β strand and an α helix (Table 4). The observed relative rigidity of the four absolutely conserved acidic residues (average S^2_{MD} = 0.90) is probably due to the hydrogen bonding contacts made by the backbone nitrogen, the preceding carbonyl oxygen

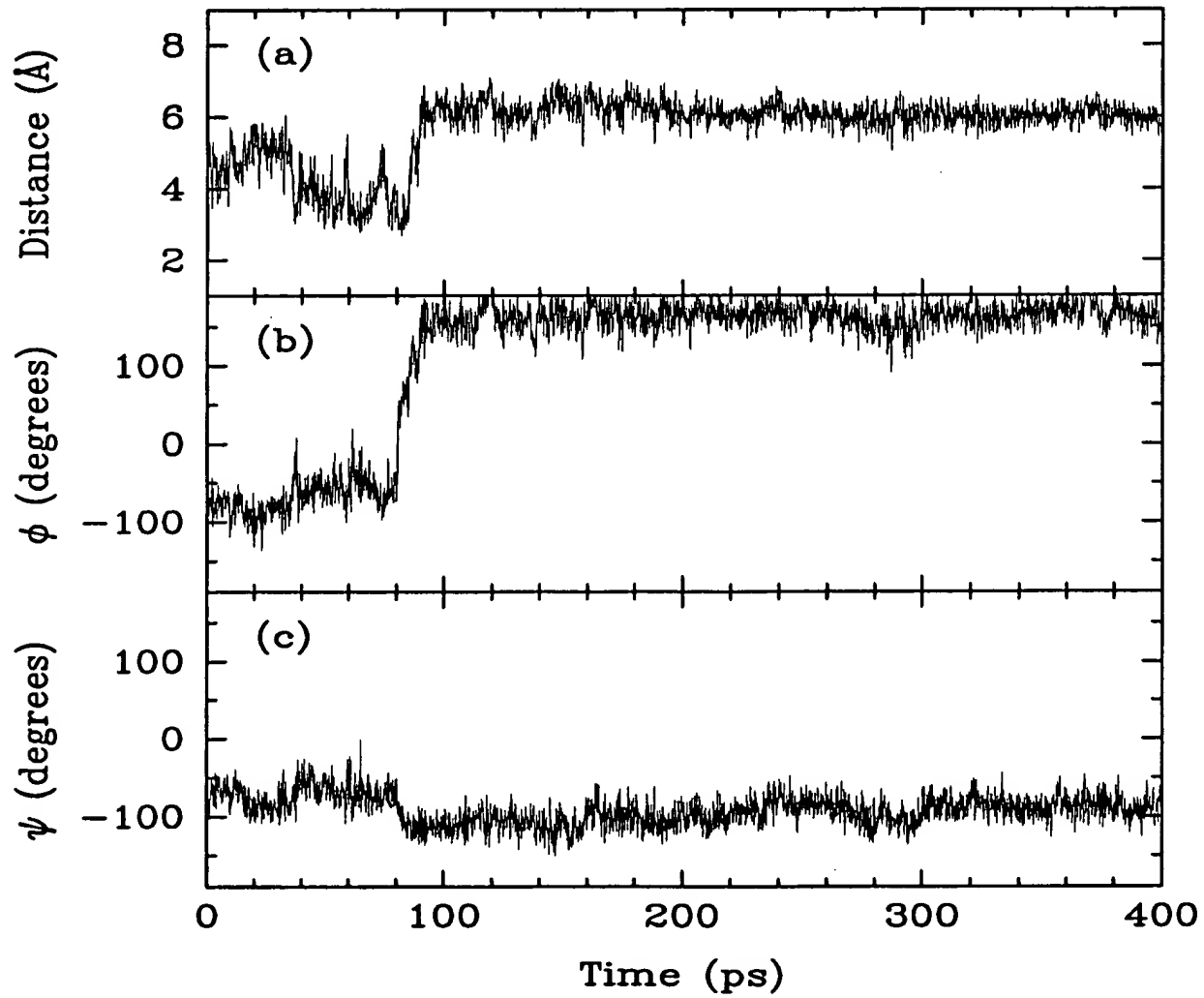

Figure 8. The distance time-series of Lys122(O)-Ala125(N) and the φ and ψ dihedral angles of Gly123 as a function of simulation time plotted in panels (a), (b) and (c), respectively.

and perhaps the side-chain carboxylate oxygens (Table 2).

Ser71

This has one of the lowest S^2_{NMR} values (0.60(±0.08)) and highest chemical exchange terms (4.6(±1.4) s^{-1}) in the protein, indicating substantial flexibility on the ps-ns and μs-ms timescale, respectively (Mandel *et al.*, 1995). On the other hand, a hydrogen bond between Ser71(N) and Thr69($O^{\gamma 1}$) is observed in both the 1RNH and 2RN2 structures (Yang *et al.*, 1990; Katayanagi *et al.*, 1992). This bond is stable during the simulation, giving rise to a S^2_{MD} value of 0.88(±0.01), consistent with the restricted motion of an N–H bond in its hydrogen bonding state. There seems to be a structural difference in the environment of Ser71 in going from the crystal to the solution state related to the side-chain conformation of Thr69. This may be caused by crystal contacts in the 2RN2 structure, as there are two intermolecular hydrogen bonds in the crystal lattice connecting neighbouring Asp70 carboxylate oxygens to Lys87(N^{ζ}) of an adjacent molecule (Katayanagi *et al.*, 1992).

His124

There are no experimental data for this catalytically important residue; its auto-correlation function shows a class III-type behaviour (Figure 3) yielding a S^2_{MD} value of 0.68(±0.02). This is midway between the S^2_{MD} values of Gly123 and Ala125, which are in agreement with the corresponding S^2_{NMR} values to within experimental uncertainty (Table 3). The flexibility of the His124 N–H bond is consistent with the finding that its backbone (and side-chain) atoms do not interact with any protein atoms, although they do form and break hydrogen bonds with several water molecules during the simulation. The backbone dihedral angles ϕ and ψ of His124 underwent no transitions during the simulation, in contrast to those of Gly123 and Ala125. However, a significant displacement of the His124 side-chain with respect to its 2RN2 coordinates occurred during equilibration (see Figure 10); this may be due to the relief of several intermolecular crystal contacts involving the backbone and side-chain nitrogens of His124.

Interestingly, the catalytically important His124 adopts different conformations with respect to the side-chains of Asp10, Glu48 and Asp70 in the crystal structures of the free enzyme (Yang *et al.*, 1990; Katayanagi *et al.*, 1992) and the active immobilized-metal-affinity-chromatography (IMAC) HIV-1 RNase H domain (Chattopadhyay *et al.*, 1993). In the 2RN2 structure (Katayanagi *et al.*, 1990) the imidazole of His124 is roughly 10 Å away from the active site residues, whereas in the 1RNH structure (Yang *et al.*, 1990), it is within 4 Å of the Asp10 and Asp70 carboxylates (see Figure 9). The conformation of the imidazole in the IMAC structure is similar to that in 2RN2, in that it is

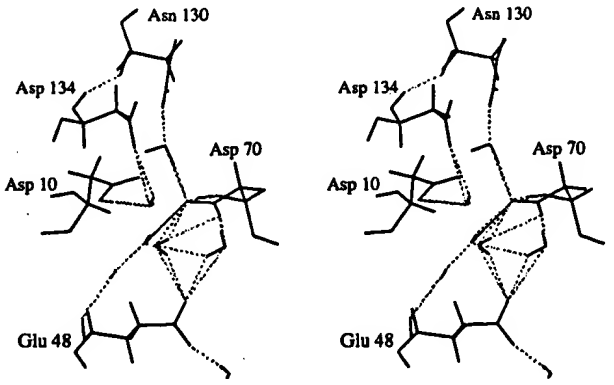

Figure 9. Hydrogen-bond network in the active site of the dynamics-averaged structure. Hydrogen bonds are represented as broken lines. For clarity, the water molecules are not labelled.

significantly displaced from the active site (Chattopadhyay *et al.*, 1993). Crystal contacts between Asp10 and Asp70 with Lys87 of a neighbouring molecule in the 2RN2 structure may be responsible for the side-chain conformation of His124. However, crystal packing is found not to be responsible for the IMAC conformation. From [^{1}H]NMR pH titration experiments (Oda *et al.*, 1993a), the pK_a value of His124 was found to be 7.1. Since this value is slightly higher than the "normal" pK_a value of an imidazole group (around 6.5 for His residues in denatured barnase), His124 was suggested to be in an acidic environment; however, as the pK_a of His124 is lower than that of His127 (7.9), whose side-chain $N^{\epsilon 2}$ hydrogen bonds to Glu119($O^{\epsilon 1}$), His124 was thought not to interact directly with the acidic side-chains (Oda *et al.*, 1993a); hence, it was concluded that it adopted a solution conformation similar to that in the 1RNH structure. The present simulation, however, suggests that another plausible cause for the observed pK_a value of His124 may be its interaction with Glu131 *via* a water molecule (His124(O)–Bul3586–Glu131($O^{\epsilon 2}$); not present in the X-ray structures) rather than its proximity to the active site acidic residues.

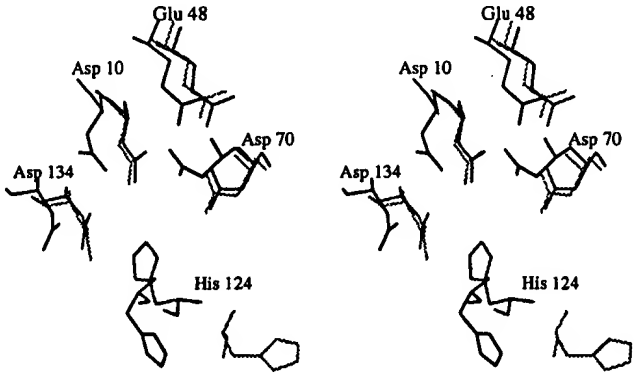

Figure 10. Conformations of the active site residues and His124 in the 2RN2 (light grey), 1RNH (dark grey) and dynamics-average structures (black).

Relevance to biological function

Experimental studies on RNase H_I have identified several regions that are important for substrate binding and/or catalytic reaction (see Introduction). The regions implicated are: (1) the active site, (2) the β1–β2 turn, (3) the basic protrusion, consisting of αC (residues 82 to 88) and the handle region (residues 89 to 99) and (4) the β5–αE loop. The possible connection between the dynamical properties of the above regions and biological function is discussed below.

The active site

In both proposed models for the catalytic mechanism of the enzyme, a water molecule is activated by either a divalent cation or a carboxylate (general base) for nucleophilic attack of the P–$O^{3'}$ bond; in a variant of the relay mechanism, His124 has been thought to enhance the catalytic rate by directly or indirectly (through a water molecule) interacting with the general base (Oda *et al.*, 1993a). A number of water molecules form stable hydrogen bonds with the four completely conserved acidic residues (see Figure 9). In some instances, these water molecules bridge two or more acidic residues together, thus shielding the repulsion of their negatively charged side-chains; Glu48($O^{\epsilon2}$)–Wat166–Asp70($O^{\delta1}$), Asp10($O^{\delta2}$)–Bul1184–Asp134($O^{\delta1}$), and Glu48($O^{\epsilon2}$)–Bul3158–Asp70($O^{\delta2}$). Asp70 is also linked to the highly conserved Asn130 *via* a water bridge: Asp70($O^{\delta1}$)–Wat175–Asn130($N^{\delta2}$). The latter water molecule (Wat175) further connects Asp70 to the neighbour of the catalytically important Asp10, i.e. Asp70($O^{\delta1}$)–Wat175–Thr9(O). In addition, two crystallographic water molecules are connected by a stable hydrogen bond: Wat166–Wat211 (Figure 9). These interactions are stable throughout the simulation (Table 2).

One or more of the structural water molecules identified above may participate in the enzymatic reaction. This is supported by the finding that replacement of Asp134 with residues other than Asn or His decreased activity significantly without serious effects on binding (Haruki *et al.*, 1994). Since the D134N and D134H mutant proteins have a carbonyl O and an imino N at the δ position similar in volume and electronegativity to that of the carboxyl O of Asp134, it is likely that D134N($O^{\delta1}$) and D134H($N^{\delta1}$) can retain the hydrogen bonding bridge with Bul1184–Asp10($O^{\delta2}$). Hence, the Asp134($O^{\delta1}$)–Bul1184–Asp10($O^{\delta2}$) interaction may be required to hold the water molecule and/or Asp10 at their appropriate positions in the active site for RNA hydrolysis. Furthermore, the low computed and observed *B*-factors, as well as the high S^2_{MD} values of Asp10, Glu48, Asp70 and Asp134, suggest that the precise alignment of these catalytically important residues may be crucial for the enzymatic reaction (Lim & Tole, 1992). Interestingly, the loss of activity of the isolated HIV-1 RNase H domain correlates with relatively low backbone S^2_{NMR} values (<0.7) of the two residues corresponding to positions 10 and 48 in the *E. coli* enzyme (the S^2_{NMR} values of the other two acidic residues corresponding to positions 70 and 134 have not been quantified; Powers *et al.*, 1992).

β1–β2 turn

Residues 14 to 17, in particular Asn16, have been implicated in substrate binding by NMR measurements (Nakamura *et al.*, 1991) as well as theoretical models (Yang *et al.*, 1990; Katayanagi *et al.*, 1992). Furthermore, the residue preceding the β1–β2 turn, Cys13, has also been implicated in binding as the introduction of a disulfide bond between Cys13 and Asn44 → Cys, and thereby the additional rigidification of the local conformation, nearly abolished enzymatic activity. Both computed and NMR S^2 values of residues Cys13–Asn16 indicate substantial mobility. The flexibility of Cys13–Asn16 probably facilitates the occurrence of conformational changes upon substrate binding. This is supported by the observation that rigid body docking of a RNA·DNA decamer onto RNase H_I resulted in a few clashes involving the β1–β2 turn, which can hence be relieved by conformational adjustment (Yang *et al.*, 1990). The loss of activity associated with the artificial 13–44 disulfide bond correlates with the lack of hydrogen bonding propensity of Cys13(S^γ): although the Cys13(S^γ)–Gly15(O) and Cys13(S^γ)–Gly18(O) distances are ⩽3.5 Å both in the 2RN2 and MD-average structures, Cys13(H^γ)–Gly15(O) and Cys13(H^γ)–Gly18(O) are greater than 2.5 Å (see Results for the definition of H-bond) in the dynamics-average structure and the 2RN2 structure with built-in hydrogens, implying the absence of hydrogen bonding interactions.

The basic protrusion

The role of residues 81 to 99 in binding DNA/RNA-hybrid and catalysis has been suggested in a model of *E. coli* RNase H_I complexed with a 21-mer DNA/RNA hybrid (Katayanagi *et al.*, 1992) and alanine-scanning mutagenesis of the basic amino acid residues (Kanaya *et al.*, 1991a). The absence of this lysine and tryptophan-rich region in the isolated HIV-1 RNase H domain has been correlated with the lack of productive substrate binding, hence its inactivity (Kanaya *et al.*, 1991a; Stahl *et al.*, 1994). However, incorporation of the basic protrusion into the equivalent position of the isolated HIV-1 RNase H domain restored activity (Stahl *et al.*, 1994), providing further evidence for the substrate-binding role of this region.

The dynamic properties of αC are obscure: unusually high *B*-factors in the simulation and the 2RN2 structure, but not in 1RNH, and S^2_{MD} and S^2_{NMR} values representative of putative secondary structure. In contrast, the handle region (residues 89 to 99) exhibits the lowest average *B*-factor among all non-regular secondary structure elements in both

1RNH and 2RN2 structures, and correspondingly, the second highest average S^2_{NMR} and S^2_{MD} values. Only Lys96 has been assigned S^2_{NMR} values for both a slow and a fast timescale (Mandel *et al.*, 1995). On the other hand, chemical exchange terms have been identified for five residues in the basic protrusion, indicating substantial mobility on the μs-ms timescale. Site-specific mutations of each of the basic residues to alanine yielded K_m values that were three to fivefold higher than that of the wild-type, except for Lys86; similar replacement of either Trp85 or 90 also resulted in greatly enhanced K_m but reduced V_{max} values, suggesting that conformational changes possibly involving disruption of the hydrogen bond connecting positions 85 and 90 (see above) affect substrate binding affinity and catalytic efficiency. These findings suggest that the concentrated positive charge of the basic protrusion and its conformation, rather than the ps-ns mobility of the handle region, may be an important determinant for binding (Kanaya *et al.*, 1991).

β5–αE loop

Residues 121 to 126 have been implicated in binding substrate from model structures of the enzyme-substrate complex (Yang *et al.*, 1990; Katayanagi *et al.*, 1992); in addition, His124 has been suggested to participate in the catalytic reaction (see Introduction). In the crystal structure of the inactive HIV-1 BH10 RNase H domain (Davies *et al.*, 1991), this loop is disordered. In contrast, in the X-ray structure of a catalytically active IMAC-RNase H domain (Chattopadhyay *et al.*, 1993), the loop is ordered, but it exists in more than one conformation. This finding is consistent with the simulation: the β5–αE loop undergoes a sharp transition, resulting in two distinct conformers (see Figures 7 and 8). Furthermore, the observed pK_a value of His124 appears to be consistent with the formation of the His124(O)-Bul3586-Glu131($O^{\epsilon 2}$) interaction during the simulation; in the latter conformation, His124 is displaced several Å from the active site carboxylates, as found in the 2RN2 and IMAC structures (Figure 9). This conformation is not in apparent contradiction to the suggested role of His124 in regulating catalysis, since the flexibility of the β5–αE loop may help to position the conserved His124 adjacent to the cluster of carboxylate side-chains upon substrate binding so that it can still take part in the enzymatic reaction.

Concluding Discussion

The present simulation of *E. coli* RNase H_I in solution yields results in overall accord with experimental results: agreement between the computed and X-ray *B*-factors is qualitative, whereas agreement between theoretical and experimental order parameters is for many residues as good as quantitative. Plausible explanations for the observed discrepancy between theory and experiment have been presented. The overestimation of the X-ray *B*-factors relative to the computed values has been deduced to be largely due to under-correction of the lattice disorder and reflections for radiation damage as well as the presence of phonon vibrations in the crystal, all of which resulted in enhanced X-ray *B*-factors. On the other hand, the overestimation of the S^2_{MD} values of most residues relative to the corresponding S^2_{NMR} values for differences $\leqslant 0.05$ has been attributed to the neglect of zero-point N–H vibrational motion. Differences between S^2_{MD} and S^2_{NMR} greater than (± 0.05), found for residues in the β1–β2 and β2–β3 turns, the coiled-coil interface between helices A and D, and the β5–αE loop, are probably due to inadequacies in the simulation (e.g. limited sampling of the present trajectory or inadequate solvation) and/or the experimental fitting protocol.

Order parameters for several highly conserved residues around the active site, whose S^2_{NMR} could not be derived due to exchange broadening or resonance overlap, have been computed; these include the catalytically important residues, Asn45, Glu48, Asp70 and His124. Although high *B*-factors have generally been correlated with low order parameters, there is no such correlation between the relatively high *B*-factors, yet high order parameters of αC were found both experimentally and in the simulation. This difference illustrates the variance in dynamical information that *B*-factors and order parameters can provide. *B*-factors are directly related to atomic rms fluctuations, whereas N–H order parameters are only sensitive to rotational motions and the local environment of the N–H bond, and may therefore reflect more localized motions as compared to *B*-factors.

The physical basis of some interesting and unusual patterns of the observed dynamics has been elucidated. The alternating motional behaviour of residues 116 to 120 in β5 and residues 32 to 36 in β3 correlates with the hydrogen bonding patterns of the backbone nitrogen atoms. The NMR-observed rigidity of the Trp90 backbone (Mandel *et al.*, 1995) is found to be due to hydrogen bonds made by Trp90(H^N) with Trp85(O) and the preceding atom Gly89(O) with Lys91(H^N), rather than the Trp90(O)–Val98(H^N) hydrogen bond in the crystal structure (Katayanagi *et al.*, 1992), which was not maintained during the simulation. The relatively significant τ_{int} values found experimentally for Gly123 and Ala125 (Mandel *et al.*, 1995) have been correlated with a concerted transition in the scalar products of the Gly123 and Ala125 N–H bonds relative to their initial orientation at about 80 ps into the production trajectory (Figure 7). An alternative basis for the measured pK_a value of His124 (7.1) is proposed: the interaction of His124(O) *via* a water bridge to Glu131($O^{\epsilon 2}$) seems consistent with its pK_a value being slightly higher than the normal pK_a value (6.5) of an imidazole group but lower than that of a histidine (≈ 8) whose side-chain is interacting directly with an acidic group.

The simulation has also provided key insights into the possible roles of several residues in the biological function of RNase H. A number of water molecules formed stable hydrogen bonds with the four absolutely conserved acidic residues in the active site, and one or more of these structural water molecules may participate in the enzymatic reaction. In particular, the Asp134($O^{\delta1}$)–Bul1184–Asp10($O^{\delta2}$) interaction may be required to hold the water molecule and/or Asp10 at their appropriate positions in the active site for RNA hydrolysis. The high S^2_{MD} values and low *B*-factors of the active site acidic residues imply that precise positioning of these residues may be crucial for the enzymatic reaction. On the other hand, the relatively low S^2_{MD} values and high *B*-factors of residues known to participate in substrate binding, Cys13, Gly15 and Asn16 in the β_1–β_2 turn, and residues 123 to 126 in the β_5–α_E loop, suggest an induced-fit mechanism for binding. In addition, the substantial flexibility of αC may play a role in positioning the unusually rigid handle region during substrate binding. Interestingly, the β5–αE loop is found to exist in two conformations, consistent with the IMAC-RNase H domain finding that this loop exhibits "discrete disorder". It appears that the flexibility of the β5–αE loop not only facilitates substrate binding, but also helps position the catalytically important His124 from a conformation distant from the active site prior to binding to an optimal one near the carboxylate side-chains once the substrate is bound.

Methods

Simulation procedure

A molecular dynamics simulation, using the program CHARMM (Brooks *et al.*, 1983), was carried out on fully solvated *E. coli* RNase H_I consisting of 2460 protein atoms and 11,484 TIP3P (Jorgensen *et al.*, 1983) water atoms, starting from the X-ray structure of Katayanagi *et al.* (1992 PDB entry 2RN2). The simulation was carried out at pH 5.5, the pH at which the NMR experiments were performed (Mandel *et al.*, 1995). The measured pK_a values of all carboxyl groups (Oda *et al.*, 1994) and histidines (Oda *et al.*, 1993a) indicate that at pH 5.5, the aspartic and glutamic acids are deprotonated except for Asp10 (pK_a = 6.1) and the histidines are protonated except for His114 (pK_a < 5.0) . The pK_a values of the other ionizable groups in RNase H_I have not been measured and are assumed to be normal: thus, all lysines and arginines are positively charged.

The forces on the atoms and their dynamics were calculated with the program CHARMM (Brooks *et al.*, 1983; version 23) and an all-hydrogen parameter set (MacKerell *et al.*, 1992). All protein and solvent atoms were propagated according to Newton's equations using a leap-frog integrator and a timestep of 1 fs. The water molecules were subjected to a deformable mean-field potential (Berkowitz & McCammon, 1982); this accounts for interactions with the infinite bulk waters that were not treated explicitly in the simulation. Electrostatic interactions were truncated by applying a force-switching function (Steinbach & Brooks, 1994) in the region between 11 and 15 Å, whereas van der Waals' interactions were truncated at 15 Å by a shifted potential (Brooks *et al.*, 1983). The non-bond list was updated every 10 fs at a 16 Å cutoff.

Initially, the 2RN2 structure was subjected to 100 steps of steepest descent and 500 steps of adopted-basis Newton-Raphson minimization in the presence of strong (10 kcal/Å^2) harmonic constraints on all heavy atoms; this relieved close contacts in the protein without disrupting its overall conformation. At this stage, the rmsd from the crystal coordinates was 0.072 Å for protein heavy atoms and 0.040 Å for backbone atoms. The protein was then immersed in a 33.3 Å radius sphere of TIP3P water molecules (Jorgensen *et al.*, 1983) constructed from an equilibrated configuration of TIP4P waters at a pressure of 1 atmosphere and a density of 0.0334 molecules/Å^3. Overlaid water molecules whose oxygens and hydrogens were within 2.8 Å and 2.0 Å, respectively, of protein heavy atoms or crystal water oxygens were deleted; the resulting system consisted of 2460 protein atoms, 225 crystal waters and 3603 overlaid bulk water molecules yielding a total of 13,944 atoms. The solvated system was then subjected to 350 steps of steepest descent minimization, during which harmonic restraints on the protein heavy atoms were gradually reduced to zero; the rmsd from the 2RN2 structure at this point was 0.203 Å for protein heavy atoms and 0.120 Å for backbone atoms.

The solvated system was then prepared for subsequent dynamics by an initial 6 ps phase of heating, during which initial velocities were randomly assigned from a Maxwell-Boltzmann distribution at 25 K and re-scaled at temperatures increasing from 25 to 300 K (for 6 ps). This was followed by a 24 ps equilibration stage, during which velocities were re-scaled to 300 K at 2 ps intervals if the average temperature fell outside a 10 K window (for 14 ps) and a 5 K window (for 10 ps). A total of 3.2 ps of production dynamics were completed per day on an IBM RS/6000 590 workstation. Co-ordinates were saved every 100 fs resulting in 4000 frames for analysis. The results described in the text are based on a 400 ps trajectory (characterized by an average temperature of 296 K), which was used to compute *B*-factors (Brooks *et al.*, 1988), internal correlation functions and order parameters (Levy *et al.*, 1981, 1982; Philippopoulos & Lim, 1994). The definitions and equations used to compute these quantities are summarized in the next section.

Analysis

The *B*-factors, B_i, were computed using the relation:

$$B_i = \frac{8\pi^2}{3}\langle \Delta r_i^2 \rangle \qquad (2)$$

where Δr_i is the atomic displacement for atom *i* calculated from the trajectory.

$C_{int}(t)$ is the correlation function describing internal motion and is given by,

$$C_{int}(t) = \langle r^{-6}(t)\rangle^{-1}\left\langle \frac{P_2[\cos\chi(t)]}{r^3(0)r^3(t)}\right\rangle \qquad (3)$$

with χ, the angle between the internuclear vectors $r(t)$ and $r(0)$, and r measured in a molecule-fixed frame. The internal correlation function can be redefined in terms of the order parameter, S^2, and the internal correlation time, τ_{int}. This is possible as the internal correlation function of

a Markovian motional process can be written as a series of exponentials (Lipari & Szabo, 1982):

$$C_{int}(t) = \sum_{i=0}^{N} a_i e^{-t/\tau_i} \quad (4)$$

A simple approximation to equation (4) that renders $C_{int}(t)$ exact at $t = 0$ and $t = \infty$ is (Lipari & Szabo, 1982):

$$C_{int}(t) = S^2 + (1 - S^2)e^{-t/\tau_{int}} \quad (5)$$

where

$$S^2 = \lim_{t \to \infty} C_{int}(t) = \frac{4\pi}{5} \langle r^{-6} \rangle^{-1} \sum_{m=-2}^{2} \left| \left\langle \frac{Y_{2m}(\theta, \phi)}{r^3} \right\rangle \right|^2 \quad (6)$$

and

$$\tau_{int} = \frac{1}{C_{int}(0) - S^2} \int_0^T (C_{int}(t) - S^2) dt \quad (7)$$

In equation (6), $Y_{2m}(\theta, \phi)$ are the second-order spherical harmonics as functions of the angular spherical coordinates of the internuclear vector in a molecule-fixed frame. In equation (7), T is the time after which $C_{int}(t) = S^2$, and $C_{int}(0)$ is the value of the correlation function at the first coordinate frame of the trajectory.

The jackknife method

The uncertainties in the order parameters and internal correlation times were computed using the jackknife procedure (Quenouille, 1956), which was introduced as a means of eliminating bias in estimating parameters. If quantity χ is computed from N data sets (e.g. coordinate frames), we wish to determine $\Delta\chi$. The algorithm is as follows:

1. Divide the N sets into k blocks (preferably 10 or more).
2. Compute χ_{-i}, the value χ from the $(k-1)$ blocks discounting block i.
3. Calculate χ_i^* from each of the k χ_{-i} values according to

$$\chi_i^* = k\chi - (k-1)\chi_{-i} \quad (8)$$

Note that χ is the weighted mean of χ_i^* and χ_{-i}.

4. $\Delta\chi$ is then given as the standard error in the mean

$$\Delta\chi = \left[\frac{\sum_{i=1}^{k} (\chi_i^* - \langle \chi_i^* \rangle)^2}{k(k-1)} \right]^{1/2} \quad (9)$$

where $\langle \chi_i^* \rangle$ denotes the mean of χ_i^* calculated over the k blocks.

For the calculation of the uncertainties of order parameters and correlation times by this procedure, the size of each of the k blocks must be chosen to be no smaller than the correlation time of the motional process of interest; the purpose of this choice is to ensure that the χ_i^* values are statistically independent from one another. In the present work, we chose $k = 10$ or a block size of 40 ps (overall length of the analysis trajectory is 400 ps), since most residues have τ_{int} values <40 ps.

The jackknife procedure differs from the commonly employed blocking technique (Allen & Tildesley, 1990):

$$\Delta\chi = \left[\frac{\sum_{i=1}^{k} (\tilde{\chi}_i - \langle \tilde{\chi}_i \rangle)^2}{k} \right]^{1/2} \quad (10)$$

where $\tilde{\chi}_i$ is the mean value of χ for each block and $\langle \tilde{\chi} \rangle$ the mean of $\tilde{\chi}_i$ over the k blocks. Although both methods reduce the correlations among neighbouring data sets, the jackknife procedure has the advantage of reducing the bias towards data sets showing greater amplitude of fluctuation, since χ_i is computed from all the blocks in the jackknife procedure rather than just block i as in the standard blocking technique. For sufficiently large k, the jackknife procedure has a further advantage of reducing, in general, the estimated error, which is a function of k^{-2} (equation (9)), whereas the uncertainty computed from the standard blocking method (equation (10)) is inversely proportional to k.

Acknowledgements

We are indebted to Professor A. Palmer for providing us with a preprint of the NMR work (Mandel *et al.*, 1995) prior to publication and many helpful suggestions including the use of the jackknife method. We also thank Drs M. Akke, B. Brooks, S.-L. Chan, W. Hendrikson, and A. Mandel for stimulating discussions. We are grateful to Professor M. Karplus for the CHARMM program. M.P. was supported by a grant from the Medical Research Council of Canada. This work was supported by the Institute of Biomedical Sciences, Academia Sinica, Taipei, Republic of China.

References

Allen, M. P. & Tildesley, D. J. (1990). *Computer Simulation of Liquids*, pp. 191–198, Oxford University Press, NY.

Beese, L. S. & Steitz, T. A. (1991). Structural basis for the 3'-5' exonuclease activity of *E. coli* DNA polymerase I: A two-metal-ion mechanism. *EMBO J.* **10**, 25–33.

Berkowitz, M. & McCammon, J. A. (1982). Molecular dynamics with stochastic boundary conditions. *Chem. Phys. Letters*, **90**, 215–217.

Brooks, B. R., Bruccoleri, R. D., Olafson, B. O., States, D. J., Swaminathan, S. & Karplus, M. (1983). CHARMM: A program for macro-molecular energy minimization and dynamics calculations. *J. Comp. Chem.* **4**, 187–217.

Brooks, C. L., III, Karplus, M. & Pettitt, B. M. (1988). *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics*, John Wiley & Sons, NY.

Bruschweiler, R. (1992). Normal modes and NMR order parameters in proteins. *J. Am. Chem. Soc.* **114**, 5341–5344.

Chandrasekhar, I., Clore, G. M., Szabo, A., Gronenborn, A. M. & Brooks, B. R. (1992). A 500 ps molecular dynamics simulation study of interleukin-1β in water. *J. Mol. Biol.* **226**, 239–250.

Chattopadhyay, D., Finzel, B. C., Munson, S. H., Evans, D. B., Sharma, S. K., Strakalaitis, N. A., Brunner, D. P., Eckenrode, F. M., Dauter, Z., Betzel, C. & Einspahr, H. M. (1993). Crystallographic analyses of an active HIV-1 ribonuclease H domain show

structural features that distinguish it from the inactive form. *Acta Crystallog. sect. D*, **49**, 423–427.

Clore, G. M., Driscoll, P. C., Wingfield, P. T. & Gronenborn, A. M. (1990). Analysis of the backbone dynamics of interleukin 1-β using 2D inverse detected heteronuclear [^{15}N-^{1}H]NMR spectroscopy. *Biochemistry*, **29**, 7387–7401.

Crouch, R. J. (1990). Ribonuclease H: from discovery to 3D structure. *New Biol.* **2**, 771–777.

Davies, J. F., Hostomska, Z., Hostomsky, Z., Jordan, S. R. & Matthews, D. A. (1991). Crystal structure of the ribonuclease H domain of HIV-1 reverse transcriptase. *Science*, **252**, 88–95.

Eriksson, M. A. L., Berglund, H., Hard, T. & Nilsson, L. (1993). A comparison of [^{15}N]NMR relaxation measurements with a molecular dynamics simulation: Backbone dynamics of the glucocorticoid receptor DNA-binding domain. *Proteins: Struct. Funct. Genet.* **17**, 375–390.

Farrow, N. E., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D. & Kay, L. E. (1994). Backbone dynamics of a free and a phosphopeptide-complexed Src homology 2 domain studied by [^{15}N]NMR relaxation. *Biochemistry*, **33**, 5984–6003.

Glover, I., Haneef, I., Pitts, J. E., Wood, S., Moss, D. S., Tickle, I. & Blundell, T. L. (1983). Conformational flexibility in a small globular hormone: X-ray analysis of avian pancreatic polypeptide at 0.98 Å resolution. *Biopolymers*, **22**, 293–304.

Haruki, M., Noguchi, E., Nakai, C., Liu, Y. Y., Oobatake, M., Itaya, M. & Kanaya, S. (1994). Investigating the role of conserved residue Asp134 in *Escherichia coli* ribonuclease H_I by site-directed random mutagenesis. *Eur. J. Biochem.* **220**, 623–631.

Hogrefe, H. H., Hegrefe, R. I., Walder, R. Y. & Walder, J. A. (1990). Kinetic analysis of *Escherichia coli* RNase H using DNA-RNA/DNA-DNA substrates. *J. Biol. Chem.* **265**, 5561–5566.

Huang, H.-W. & Cowan, J. A. (1994). Metallobiochemistry of the magnesium ion. Characterization of the essential metal-binding site in *Escherichia coli* ribonuclease H. *Eur. J. Biochem.* **198**, 437–440.

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. (1983). Comparison of simple potential functions for simulating water. *J. Chem. Phys.* **79**, 926–935.

Kabsch, W. (1976). A solution for the best rotation to relate two sets of vectors. *Acta Crystallog. sect. A*, **32**, 922–923.

Kanaya, S., Kohara, A., Miyagawa, M., Matsuzaki, T., Morikawa, K. & Ikehara, M. (1989). Overproduction and preliminary crystallographic study of ribonuclease H from *Escherichia coli*. *J. Biol. Chem.* **264**, 11546–11549.

Kanaya, S., Kohara, A., Miura, Y., Sekiguchi, A., Iwai, S., Inoue, H., Ohtsuka, E. & Ikehara, M. (1990). Identification of the amino acid residues involved in an active site of *Escherichia coli* ribonuclease H by site-directed mutagenesis. *J. Biol. Chem.* **265**, 4615–4621.

Kanaya, S., Katsuda-Nakai, C. & Ikehara, M. (1991a). Importance of the positive charge cluster in *E. coli* ribonuclease H_I for the effective binding of the substrate. *J. Biol. Chem.* **266**, 11621–11627.

Kanaya, S., Katayanagi, K., Morikawa, K., Inoue, H., Ohtsuka, E. & Ikehara, M. (1991b). Effect of mutagenesis of each of 5 histidine residues on enzymatic activity and stability of ribonuclease H from *Escherichia coli*. *Eur. J. Biochem.* **198**, 437–440.

Katayanagi, K., Miyagawa, M., Matsushima, M., Ishikawa, S., Kanaya, S., Ikehara, M., Matsuzaki, T. & Morikawa, K. (1990). Three-dimensional structure of ribonuclease H from *E. coli*. *Nature*, **347**, 306–309.

Katayanagi, K., Miyagawa, M., Matsushima, M., Ishikawa, S., Kanaya, S., Nakamura, H., Ikehara, M., Matsuzaki, T. & Morikawa, K. (1992). Structural details of ribonuclease H from *Escherichia coli* as refined to an atomic resolution. *J. Mol. Biol.* **223**, 1029–1052.

Katayanagi, K., Ishikawa, M. & Morikawa, K. (1993a). Crystal structure of *Escherichia coli* ribonuclease H_I in complex with Mg^{2+} at 2.8 Å resolution: proof for a single Mg^{2+}-binding site. *Proteins: Struct. Funct. Genet.* **17**, 337–346.

Katayanagi, K., Ishikawa, S., Okumura, M., Ariyoshi, M., Kanaya, S., Kawano, Y., Suzuki, M., Tanaka, I. & Morikawa, K. (1993b). Crystal structures of ribonuclease H_I active site mutants from *Escherichia coli*. *J. Biol. Chem.* **2683**, 22092–22099.

Levy, R. M., Karplus, M. & Wolynes, P. G. (1981). NMR relaxation parameters in molecules with internal motion: exact Langevin trajectory results compared with simplified relaxation models. *J. Am. Chem. Soc.* **103**, 5998–6011.

Levy, R. M., Dobson, C. M. & Karplus, M. (1982). Dipolar NMR relaxation of nonprotonated aromatic carbons in proteins: structural and dynamical effects. *Biophys. J.* **39**, 107–113.

Lim, C. & Tole, P. (1992). Endocyclic and exocyclic cleavage of phosphorane monoanion: a detailed mechanism of the RNase A transphosphorylation step. *J. Am. Chem. Soc.* **114**, 7245–7252.

Lipari, G. & Szabo, A. (1982). Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.* **104**, 4546–4559.

Luo, G., Sharmeen, L. & Taylor, J. (1990). Specificities involved in the initiation of retroviral plus-strand DNA. *J. Virol.* **64**, 592.

MacKerell, A. D., Jr, Bashford, D., Bellot, M., Dunbrack, R. L., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Wiorkiewicz-Kuczera, J. & Karplus, M. (1992). Self-consistent parameterization of biomolecules for molecular modeling and condensed phase simulations. *FASEB J.* **6**.

Mandel, A. M., Akke, M. & Palmer, A. G. III (1995). Backbone dynamics of *Escherichia coli* ribonuclease H_I: Correlations with structure and function in an active enzyme. *J. Mol. Biol.* **246**, 144–163.

Nakamura, H., Oda, Y., Iwai, S., Inoue, H., Ohtsuka, E., Kanaya, S., Kimura, S., Katsuda, C., Katayanagi, K., Morikawa, K., Miyashiro, H. & Ikehara, M. (1991). How does RNase H recognize a DNA-RNA hybrid? *Proc. Natl Acad. Sci. USA*, **88**, 11535–11539.

Oda, Y., Nakamura, H., Kanaya, S. & Ikehara, M. (1991). Binding of metal ions to *E. coli* RNase H_I observed by [^{1}H-^{15}N] heteronuclear 2D NMR. *J. Biomol. NMR*, **1**, 247–255.

Oda, Y., Yoshida, M. & Kanaya, S. (1993a). Role of histidine 124 in the catalytic function of ribonuclease H_I from *Escherichia coli*. *J. Biol. Chem.* **268**, 88–92.

Oda, Y., Iwai, S., Ohtsuka, M., Ishikawa, E., Ikehara, M. & Nakamura, H. (1993b). Binding of nucleic acids to

E. coli RNase H_I observed by NMR and CD spectroscopy. *Nucl. Acids Res.* **21**, 4690–4695.
Oda, Y., Yamazaki, T., Nagayama, K., Kanaya, S., Kuroda, Y. & Nakamura, H. (1994). Individual ionization constants of all the carboxyl groups in ribonuclease H_I from *Escherichia coli* determined by NMR. *Biochemistry*, **33**, 5275–5284.
Palmer, A. G. III & Case, D. A. (1992). Molecular dynamics analysis of NMR relaxation in a zinc-finger peptide. *J. Am. Chem. Soc.* **114**, 9059–9067.
Petsko, G. A. (1975). Protein crystallography at sub-zero temperatures: Cryo-protective mother liquors for protein crystals. *J. Mol. Biol.* **96**, 381–392.
Petsko, G. A. & Ringe, D. (1984). Fluctuations in protein structure from X-ray diffraction. *Annu. Rev. Biophys. Bioeng.* **13**, 331–371.
Philippopoulos, M. & Lim, C. (1994). Internal motions in the molecular tumbling regime: Effect on NMR dipolar cross-relaxation and interproton distance determination. *J. Phys. Chem.* **98**, 8264–8273.
Post, C. B., Dobson, C. M. & Karplus, M. (1989). A molecular dynamics analysis of protein structural elements. *Proteins: Struct. Funct. Genet.* **5**, 337–354.
Powers, R., Clore, G. M., Stahl, S. J., Wingfield, P. T. & Gronenborn, A. (1992). Analysis of the backbone dynamics of the ribonuclease H domain of the human immunodeficiency virus reverse transcriptase using 15N relaxation measurements. *Biochemistry*, **31**, 9150–9157.
Quenouille, M. H. (1956). Notes on bias in estimation. *Biometrika*, **43**, 353–360.
Smith, P. E., van Schaik, R. C., Szyperski, T., Wuthrich, K. & van Gunsteren, W. F. (1995). Internal mobility of the basic pancreatic trypsin inhibitor in solution: A comparison of NMR spin relaxation measurements and molecular dynamics simulations. *J. Mol. Biol.* **246**, 356–365.
Stahl, S. J., Kaufman, J. D., Vikic-Topic, S., Crouch, R. J. & Wingfield, P. T. (1994). Construction of an enzymatically active ribonuclease H domain of human immunodeficiency virus type 1 reverse transcriptase. *Protein Eng.* **7**, 1103–1108.
Steinbach, P. J. & Brooks, B. R. (1994). New spherical-cut-off methods for long-range forces in macromolecular simulation. *J. Comp. Chem.* **15**, 667–683.
Stone, M. J., Fairbrother, W. J., Palmer, A. G. III, Reizer, J., Seier, M. H. Jr & Wright, P. E. (1992). Backbone dynamics of the *Bacillus subtilis* glucose permease IIa domain determined from [^{15}N] NMR relaxation measurements. *Biochemistry*, **31**, 4394–4406.
Teeter, M. M. & Hendrickson, W. A. (1979). Highly ordered crystals of the plant seed protein crambin. *J. Mol. Biol.* **127**, 219–224.
Uchiyama, Y., Miura, Y., Inoue, H., Ohtsuka, E., Ueno, Y., Ikehara, M. & Iwai, S. (1994). Studies of the interactions between *E. coli* RNase H_I and its substrate. *J. Mol. Biol.* **243**, 782–791.
van Gunsteren, W. F. & Berendsen, H. J. C. (1977). Algorithms for macromolecular dynamics and constraint dynamics. *Mol. Phys.* **34**, 1311–1327.
Yang, W., Hendrickson, W. A., Crouch, R. J. & Satow, Y. (1990). Structure of ribonuclease H phased at 2 Å resolution by MAD analysis of the selenomethionyl protein. *Science*, **249**, 1398–1405.

Edited by B. Honig

(Received 7 June 1995; accepted 18 September 1995)